United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Development Division
Washington, DC 20250

RDD Research Report
Number RDD-02-01

June 2002

Sample Allocation for Estimation of the Number of "Not on Mail List" (NML) Farms for the 2002 Census of Agriculture

Raj S. Chhikara
Floyd M. Spears
Charles R. Perry

United States
Department of
Agriculture

National Agricultural Library

EXECUTIVE SUMMARY

The goal of the Census of Agriculture is to collect data for the population of farms. The primary tool for collecting data is a questionnaire mailed to farm operations which are on the National Agricultural Statistics Service's (NASS) mail list. Evaluation of the 1997 Census of Agriculture results shows that 13.7 percent of the total number of farms containing 2.2 percent of the acreage in farms were "not on mail list" (NML) operations. Since the mail list is an incomplete frame, it is necessary to use sampling methods for making estimates for the NML farms. This report describes, evaluates, and recommends a method for adjusting the allocation of NASS's area frame sample to achieve an NML estimates with the desired reliability.

The current NASS area frame sample design and allocation is quite efficient for most major crops. It is designed to support probability-based sample surveys conducted for major crops, livestock, economics, and environmental data. It is, however, not very efficient for most rare items. The number of NML farms and their acreage are not reliably estimated with the current area frame allocation since they are rare items. Changes in the sample allocation are necessary to improve the precision of their estimates for Census of Agriculture years.

This study shows that the desired reliability can be achieved at a minimal cost by starting with the current sample allocation and judiciously adding sample segments where required to achieve a better estimate of the NML farm count. The new samples are determined by including the stratum variances for NML farm count as input in the multivariate sample allocation procedure used by NASS. It is approached two ways. One approach uses design-based estimates of the stratum variances for NML farm counts. The design-based estimates are obtained using the area frame closed estimates and "pseudo not on the list" (pseudo NOL) estimates derived from the 1999 June Agricultural Survey (JAS) data. The pseudo NOL estimates are derived using all 1999 JAS farm observations that are not on any of the NASS national survey lists during 1999. The other approach uses model-based estimates of the stratum variances for NML farm count, which are expected to be more stable than the design-based estimates. The model-based estimates are obtained by modeling the pseudo NOL variance estimates in terms of the total number of farms. The total number of farms is more reliably estimated than the number of pseudo NOL farms in a stratum.

In modeling the stratum variance for the number of pseudo NOL farms, variability is investigated with respect to the level of aggregation and the possible factors of heterogeneity. The occurrence of pseudo NOL farms is assumed to behave like a Poisson process. This assumption is extended to relate the stratum variance to the total number of farms where the relationship is linear using a logarithmic scale in the predictor and the response variable. The heterogeneity in variance is mostly accounted for by the land use strata embedded into the broader category strata numbers as 10-19, 20-29, 30-39, 40-49, and 50-59. This evaluation of variances at the sampling substratum level leads to the consideration of a mixed linear model for the substratum variance. The proposed model (given by Equation 3 on page 9) is judged to be appropriate for stratum variance estimation.

The census district is determined to be the most suitable level for fitting the model to data. The weighted general linear model using the number of sample segments as the weight, provides significant data fits with R-squares for the model fits ranging from 0.321 to 0.746 across the eleven census districts. In all cases the model fits are highly significant ($p < 0.001$).

The number of NML farms is added to the eight items currently used by NASS for determining the optimal sample allocation for a June Area Frame Survey. The allocation procedure is carried out separately using both the design-based and the model-based stratum variance estimates. The allocation is determined in each case so that the total number of sample segments would be approximately 10 percent more than in the 1999 JAS, resulting in a reduced CV (coefficient of variation) for the NML farm counts for the 2002 Census of Agriculture. Each of these sample allocations is further modified to compensate for the smaller sample sizes in many strata. The new allocation is compared with the 2001 sample allocation, and the larger of these two sample sizes for each stratum is recommended for the 2002 area frame sampling survey. The final recommended allocation has approximately 15 percent more sample segments than the 2001 sample allocation.

The precision achievable under the final allocation for the NML farm count estimate is evaluated at the state and national levels. The desired level of precision is a "board CV" (which is defined as the standard error of the NML farm count estimate divided by the Agricultural Statistics Board estimate) of 6.5 percent at the state level. States not achieving that level are combined with other similar adjoining states and, if necessary, samples are added to achieve this level of precision. This results in an achievable board CV of 0.5 percent at the national level under the recommended 2002 area frame sample allocation.

RECOMMENDATIONS

1. The area frame sample allocation given in Tables A2 and A3 of this report, which includes the supplemental sample allocation derived from the model-based variance estimates developed in this study, is recommended for FY 2002 to support estimation of the number of NML farms in the 2002 Ag Census.

2. The multivariate area frame sample allocation procedure should use stratum variance estimates based on mixed model procedures for all agricultural items. This will result in a more reliable and efficient allocation. In particular, mixed model stratum variance estimation procedures based on generalizations of the variance models employed by Perry (1992) should be investigated for derivation of stratum variance estimates for all crop and livestock items for use in area frame sample allocation.

Sample Allocation for Estimation of the Number of "Not on Mail List" (NML) Farms for the 2002 Census of Agriculture

Raj S. Chhikara, Floyd M. Spears, and Charles R. Perry

A mixed model approach was utilized to develop stratum variance estimates for use in determining the sample allocation for estimation of "Not on Mail List" (NML) farm counts in the 2002 U.S. Census of Agriculture. Sample allocation was carried out by adding the NML farm item to the set of eight items currently used in the determination of sample allocation for the June Area Frame Survey. Changes in the sample allocations are recommended in the area frame sample design to support both the current sample survey based estimates and the 2002 Census of Agriculture NML farm count estimates.

KEY WORDS: Census of Agriculture; Area frame sampling; Mixed model; Sample allocation; Variance function; Pseudo NOL.

1 INTRODUCTION

The National Agricultural Statistics Service (NASS) is responsible for conducting the 2002 U.S. Census of Agriculture (Ag Census). Although NASS assumed responsibility from the Bureau of the Census for the 1997 Ag Census, much of its work was completed based on planning and processes inherited from the Census Bureau. NASS is committed to consolidating its experience and resources to improve efficiency in estimation for the 2002 Ag Census.

In the 1997 Ag Census, 13.7 percent of the total number of farms were estimated to be "Not on Mail List" (NML) operations. (For Ag Census purposes, a farm is defined to be any place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the census year.) In terms of farm land, 2.2 percent of land acreage was estimated to be due to NML farms. The percentage of farms missed in the 1997 Ag Census varies considerably by farm value of sales and by farm size in acreage.

There was a relatively high variability associated with undercoverage across farm categories in the 1997 Ag Census. This requires NASS to develop a more efficient sampling and estimation methodology for use in making undercoverage adjustments to the 2002 Ag Census results from the census list of farm operations.

An investigation of 1997 coverage estimates by Allen (1999) indicates that the occurrence of NML operations is due to a diverse set of factors. Two major factors already mentioned that influence inclusion of a farm on the census mailing list are:

Raj S. Chhikara is Professor of Statistics, University of Houston-Clear Lake, 2700 Bay Area Blvd., Houston, TX 77058. Floyd M. Spears is Associate Professor of Mathematics, Harding University, Box 10764, Searcy, AR 72143. Charles R. Perry is Senior Research Scientist, National Agricultural Statistics Service, 3251 Old Lee Highway, Room 305, Fairfax, VA 22030. The authors would like to thank Ron Bosecker, Carol House, and George Hanuschak for their management support. The feedback, comments and suggestions from Roberta Pense and Phil Kott are appreciated. The authors would also like to thank Mike Bellow for preparing the SAS data file for this project, and Suzan Benz for providing the sample allocation program and documention.

- Size of the Farm in Acreage
- Farm Value of Sales.

Table 1 summarizes the percentage of NML farms by Farm Size in Acreage. Approximately one quarter of all small farms (size less than 50 acres) are NML farms. This is a substantially higher percentage than for larger farms.

Table 1: Contribution of NML to 1997 Ag Census by Farm Size in Acreage

Farm Size in Acreage	Census + NML	Percent NML
< 10	207,050	25.7
10-49	539,608	24.8
50-179	673,298	12.6
180+	786,476	3.8
All Farms	2,206,432	13.7

Table 2 lists the percentages of NML farms by Farm Value of Sales. Here again, the smallest category in terms of Farm Value of Sales ($1,000-$2,499) has a substantially higher percentage of NML farms than the other categories.

Table 2: Contribution of NML to 1997 Ag Census by Farm Value of Sales

Farm Value of Sales	Census + NML	Percent NML
$1,000 - 2,499	689,512	29.1
$2,500 - 9,999	529,268	12.0
$10,000+	987,652	3.6

Farm characteristics that seem to be associated with NML operations are:

- Horses on farm
- Chicken Farm
- Individual/Family Farm
- Farm operator's place of residence, principal occupation, tenure and length of farm operation

Other influential factors seem to be the age, gender, and race of the farm operator. It was shown by Allen (1999) that a NML operator was more likely to be: female than male, black than white, and in the age group of 35-44 than in other age groups.

The current land use stratification employed by NASS, even though fairly detailed in terms of agricultural features, is designed to achieve reliable estimates for major agricultural crops. The area frame sampling is based on an extensive layout with delineation of primary sampling units (PSU), sampling strata (substrata), and strata across all agricultural areas in the country. Sample allocation, which is optimized for major crop acreage estimation, is made at the stratum level and proportionally allocated to the substratum level.

The current area frame sample design and allocation is quite efficient for most major crops and, like most area frames, not very efficient for most rare items. The estimates of number of NML farms and their acreage have rather low precision. Changes in the sample allocation to the NASS area frame strata are required in order to obtain more reliable coverage estimates. The most desirable approach is to use the same sample design, adding judiciously a set of sample segments to the existing sample allocation to achieve the desired precision while minimizing the increase in survey cost. Thus a major objective of this study is to determine a sample allocation for the area frame under the current stratified sample design that will support reliable NML farm count estimates.

2 DATA

The area frame closed estimates (full area frame) and pseudo "not on the list" (pseudo

NOL) component of estimates of number of farms from the 1999 June Agricultural Survey (JAS) data and the associated variances are computed at the substratum level and aggregated to the land-use stratum, state, census district, and national level for the purposes of this study. The estimates of the number of farms are computed for each combination of the various levels of the Farm Value of Sales and Farm Type factors, which were deemed potentially important for modeling the variability for the number of NML farms.

The full area frame includes observations that are on the list as well as those not on the list. NASS maintains one list frame, but subsets it for various surveys; so not all of the list is sampled in every survey. Therefore, a farm observation can be on the list of one survey and NOL for another survey. A pseudo NOL includes only farm observations that are not on *any* NASS list: thus a pseudo NOL estimate can be smaller than the regular NOL estimate.

In this study, the pseudo NOL component consists of only those farms that are not on any of the survey lists used for crops, burley tobacco, cattle, equine, ARMS, hogs, labor, peanuts, and sheep. This is viewed to be comparable to the case of NML operations for an Ag Census. However, in 1999 the various list frames used throughout the year did not cover all the records used in the Ag Census. This means area frame farms that are not on the Ag Census list (NML) would be expected to be smaller than our pseudo NOL. This difference has probably disappeared for 2002, because if we take all the list frame records that are in at least one survey during the year, we probably have essentially all the records that will be used in the 2002 Ag Census. Nevertheless, for the present analyses we are dealing with the 1999 list, which means that the pseudo NOL is somewhat larger than the actual Ag Census NML.

Table 3 lists the variables included in the study data along with a brief description of each variable. The DOMAIN variable indicates whether an observation is from the full area frame (DOMAIN=1) or the pseudo NOL (DOMAIN=2). The VALUE variable indicates whether or not an observation is from a specific farm value of sales category as defined in the table. The TYPE variable indicates the type of farm an observation is from, where the four different farm types are defined as follows:

1. Group 1: grains, oilseeds, tobacco, and cotton.
2. Group 2: vegetables, fruits, tree nuts, berries, hay, peanuts, sugar, beets, etc.
3. Group 3: beef cattle, dairy, hogs, sheep, and goats.
4. Group 4: horses, poultry, aquaculture, bees, etc.

3 MODELING STRATUM VARIANCES

In multipurpose surveys, the optimum sample allocation for estimating multiple items can be found by establishing linear variance constraints for the item estimates and solving for the allocation which minimizes survey cost (Bethel, 1986 and 1989). Variance estimates for an item of interest (such as NML farm count) can be incorporated into this multivariate procedure to arrive at the optimal sample allocation for estimating items of interest. A current implementation of the procedure by NASS utilizes land-use stratum level variance estimates for each item. Stratum variances are obtained by aggregating the substratum level variances, which are often based on a relatively small number of observations.

This paper describes a modeling approach for improving the reliability of estimates of stratum variances. The sample size varies greatly among strata, so strata with similar characteristics were collapsed (grouped together) into a single super stratum for the purpose of modeling. NASS land-use strata are organized into

Table 3: Variables in Study Data

Variable	Description
NPOP	(N_i) number of population units
NSEG	(n_i) number of sample units
STATE	state FIPS code
STRATUM	land-use stratum
SUBSTRAT	sampling stratum
T	estimate of number of farms
V	estimated variance of T
SD	estimated standard deviation of T
DOMAIN	1=full area frame 2=pseudo NOL
VALUE	0=all farms 1=\$1,000-\$2,499 2=\$2,500-\$9,999 3=$\geq$ \$10,000
TYPE	0=all farm types 1=Major Crops 2=Minor Crops 3=Major Livestock 4=Minor Livestock

groups of strata that have similar characteristics. The groupings defined in Table 4 were used to form the collapsed strata.

In general, strata are formed by percentage of cultivated land, with strata from 10-19 having the highest percentage of cultivated land. The percentage of cultivation in strata from 20-29 is lower than in strata from 10-19. The percentage of cultivation in strata from 40-49 is lower than in strata from 20-29. Strata in 30-39 are used for urban areas, which typically have a very low percentage of cultivated land. Strata 50 and above contain non-agricultural land. Although land-use strata are determined at the state level, they are similar in certain agricultural characteristics among states within an agricultural census district as defined in Table 5. For modeling purposes, the collapsed strata were considered (across states) at the census district level. No modeling was done at the state level or with state as a factor.

Table 4: Collapsed Strata

Collapsed Strata	Land-use Strata	Description
1	10 - 19	Intense Cultivation
2	20 - 29	Moderate Cultivation
3	30 - 39	Urban Area
4	40 - 49	Low Cultivation
5	≥ 50	Non-Agricultural

Table 5: Agricultural Census Districts

Census District	List of States
1	CT DE MA MD ME NH NJ NY PA RI VT WV
2	IA IL IN KS NE OH
3	MI WI
4	AL GA KY NC SC TN VA
5	FL
6	AR LA MO MS OK
7	TX
8	MN MT ND SD
9	CO NM NV UT WY
10	ID OR WA
11	AZ CA

It should be pointed out that if the current survey variance estimates are used to adjust or supplement an existing area frame allocation without completely redrawing the entire sample, a downward bias is likely to be introduced into survey estimates. This follows because most NASS area frame items of interest have right skewed distributions with a relatively large number of large observations in the right tails, which results in a positive correlation between the item means or totals and the corresponding variances. Since a supplemental allocation will add samples to some strata, particularly those with high observed variances, it will result in smaller variances accompanied by smaller means for these strata when the sample from the new survey is utilized. This in turn will lead to a lower estimate of the survey item total. Therefore, care should be exercised when adding extra segments to the current allocation in an attempt to lower the overall NOL variances.

One way to avoid any potential downward bias in the NOL estimates is to base all allocation decisions on historical data which do not come from any of the segments in current use – the requirement is that all allocation decisions must be based on data that are unrelated to data associated with the set of segments that are being used to derive the estimates of interest. Another way to avoid, or at least greatly diminish, any potential downward bias in the NOL estimates is to base all allocation decisions on model based variance estimates which tend to smooth out the effects associated with unusual large observations over a group of strata which eliminates any stratum level biases.

3.1 Variance Function

The number of NML farms in a stratum can be assumed to follow a **counting process** as a basis for modeling stratum variance. For example, if the binomial process is assumed, the variance (σ^2) is proportional to $p(1-p)$, where p is the proportion of NML farms; and if the Poisson process is assumed, then $\sigma^2 = \mu$, where μ is the mean number of NML farms. In either case, the variance is a function of the expected number of NML farms. A more general function that can be used for modeling the stratum variance is:

$$\sigma^2 = \alpha \mu_x^\beta \tag{1}$$

where α and β are constant parameters and μ_x is the mean of variable X which is related to the NML farm count and thus is being considered as a predictor for NML variance. In Equation 1, α is the proportionality constant which depends upon the stratum size and β is the power value which depends upon the predictor. This variance function has proven to be very appropriate in planning studies on crop acreage surveys as seen in Smith (1936), Mahalonobis (1946 and 1968), Perry and Hallum (1979), Chhikara and Perry (1980 and 1986) and Perry (1992 and 1994), among others. Using the logarithmic scale, the relationship in Equation 1 can be linearized as follows:

$$\ln \sigma^2 = \ln \alpha + \beta \ln \mu_x. \tag{2}$$

In the 1999 JAS data prepared for this study, the potential variance predictors that were considered are Number of Pseudo NOL Farms, Number of Total Farms, Farm Value of Sale, and Farm Type. If the counting process approximately holds, the number of pseudo NOL farms is expected to be well correlated with the variance. This is explored by plotting the substratum variance versus the number of pseudo NOL farms in the substratum. Figures 1 through 3 show a selection of typical scatter plots of variance for collapsed strata 1 and 4 of census district 6, collapsed stratum 2 of census district 4 and collapsed stratum 3 of census district 2.

In Figure 1, the design-based substratum variance is plotted against the number of pseudo NOL farms. In Figure 2, the same observations are plotted using the logarithmic transformation for both variables. Although the plots in Figure 2 show a good linear relationship, the number of pseudo NOL farms cannot be used as a predictor since it is itself to be estimated. Instead, the total number of farms is more reliably determinable and hence is a viable predictor. The logarithm of design-based substratum

Figure 1: Variance of Number of Pseudo NOL Farms versus Number of Pseudo NOL Farms

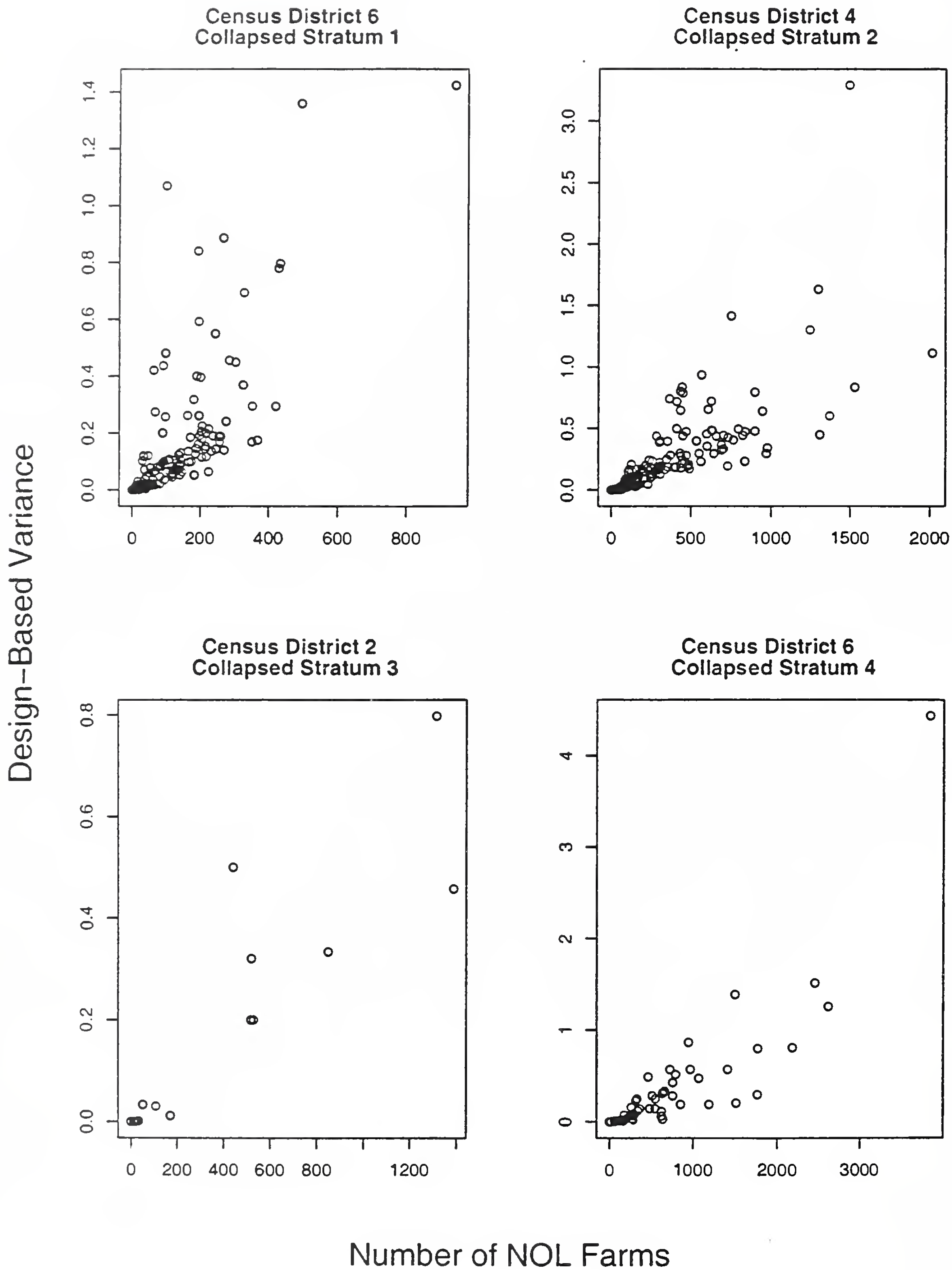


Figure 2: Ln(Variance of Number of Pseudo NOL Farms) versus Ln(Number of Pseudo NOL Farms)

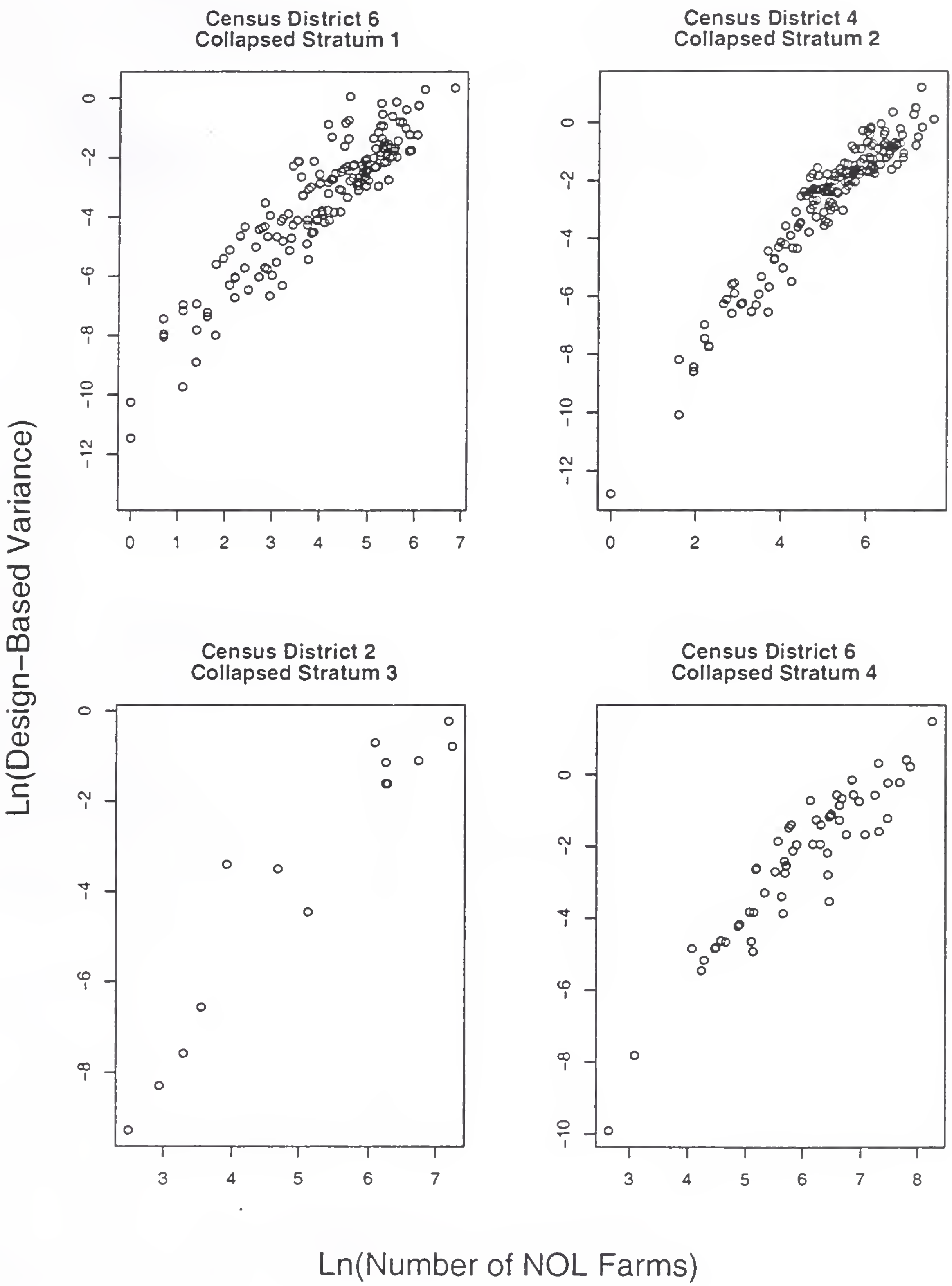
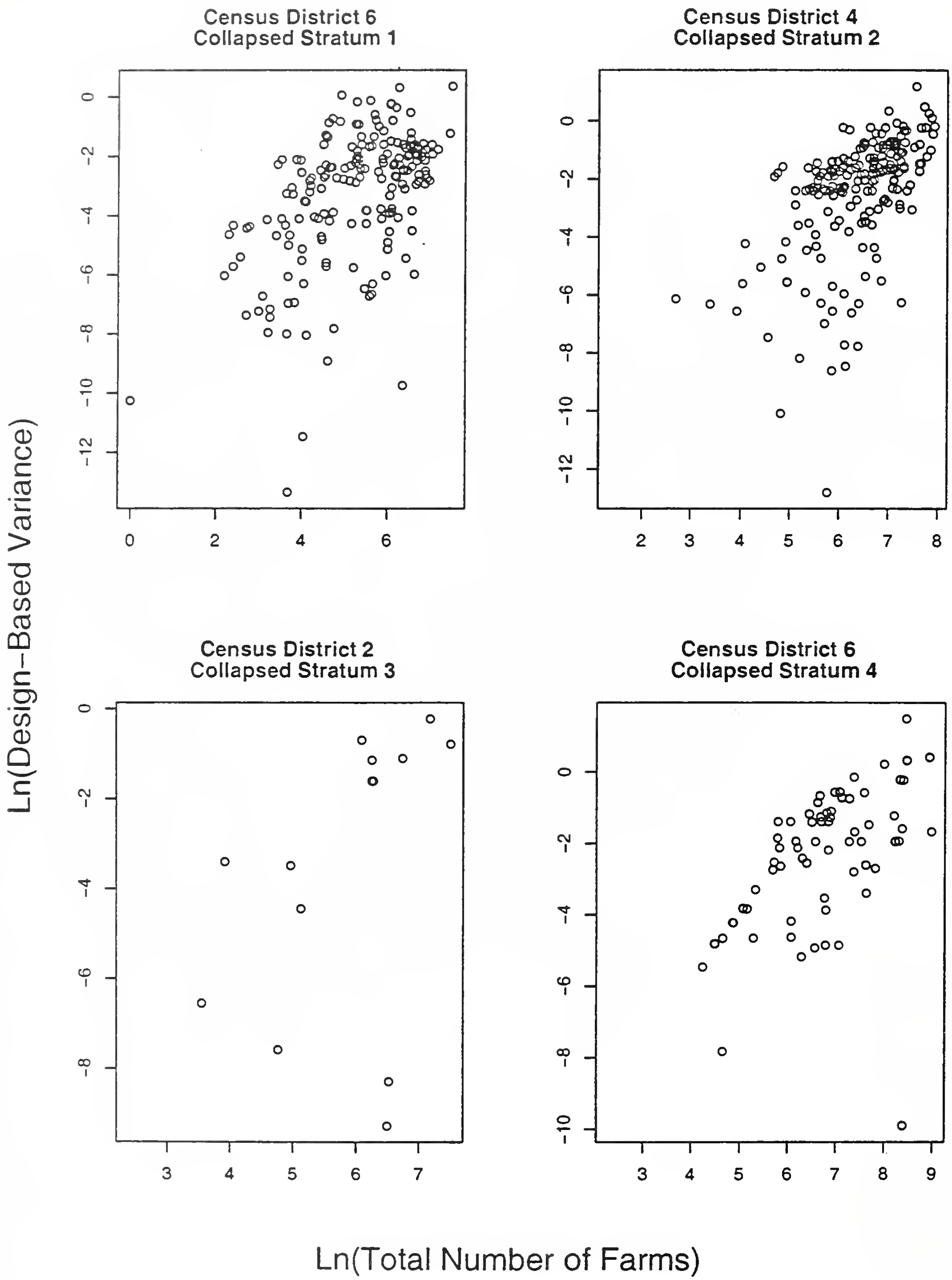


Figure 3: Ln(Variance of Number of Pseudo NOL Farms) versus Ln(Total Number of Farms)



variance is plotted against the logarithm of total number of farms in Figure 3. The total number of farms does have a linear relationship to the pseudo NOL farm substratum variance. As expected. it is not as strong a predictor of substratum variance as is the pseudo NOL number of farms.

Other predictors such as Farm Value of Sales and Farm Type were investigated by including them as factors in the general linear model. These factors were not determined to be sufficiently useful in this study to merit inclusion in the modeling of variance.

3.2 Mixed Model

The strata within a collapsed stratum have certain agricultural characteristics of their own. Thus their variances are subject to heterogeneity. which can be accounted for by introducing a **random effect** factor into the model for stratum variance. The data can be modeled at the level of collapsed strata or census district. Since the census districts are created based on similarities in broad agricultural characteristics, it is deemed appropriate to consider modeling at the census district level with the collapsed strata in a census district as the domains. This leads to a mixed model with a fixed effect due to collapsed strata and a random effect due to strata within a collapsed stratum. The model form is:

$$\ln s_{ijk}^2 = \ln \alpha + \beta \ln T_{ijk} + \tau_i + t_{j(i)} + \epsilon_{ijk} \quad (3)$$

where s_{ijk}^2 is the substratum variance estimate from the 1999 JAS data, T_{ijk} is the total number of farms, τ_i is the fixed effect due to collapsed strata, $t_{j(i)}$ is the random effect due to strata within a collapsed stratum, and ϵ_{ijk} is the error term. Note that i indexes the collapsed strata, j indexes the strata within a collapsed stratum and k indexes the substrata.

Initially, farm value of sales was included in the model as a block effect and found to be a significant factor in model fit. However, it was later removed from the model because it made little difference in the resulting sample allocation.

3.3 Model Fits

The model fit for each census district is made using NSEG (number of sample segments) as the weight for each observation. This use of weighting improves the model fit. Since the stratum variance estimates are based upon the number of sample segments, a higher value of NSEG influences the model fit more than does a smaller value.

The predicted variance estimates resulting from the model fit are adjusted to correct for the bias which results when the model predictions are transformed from the logarithmic scale to the original scale. Each predicted stratum variance is multiplied by the corresponding census district level ratio of design-based variance to model-based variance.

The model fits for each census district are summarized in Table 6 which includes estimates of the parameters $\ln \alpha$ and β, p-values for the fixed (collapsed strata) and random (strata within collapsed strata) effects, the model R-square, and a list of the strata within each collapsed stratum found to be significant. Both $\ln \alpha$ and β are significant in all of the model fits across census districts and the random effect (due to strata in collapsed strata) is significant in all census districts except for census districts 3, 4, and 11 (11 is marginal). This confirms that the land-use stratification leads to more homogeneous strata than that based on broad land agricultural features reflected in collapsed strata. Overall, each model fit is highly significant (p-value < 0.01). The variability in the substratum variances accounted for by the model ranged substantially, with R-squares from 0.321 to 0.746.

Table 6: Model Fits

Census District	Number of Observations	Estimates $\ln \alpha$	β	p-values τ_i	$t_{j(i)}$	Significant Strata	Model R-square
1	130	-11.739	0.859	0.1835	0.0390	40,42	0.321
2	207	-15.871	1.908	0.0232	0.0001	11,12,31	0.381
3	49	-14.980	1.779	0.4426	0.6301		0.359
4	232	-11.367	1.361	0.7088	0.2955		0.408
5	30	-19.144	2.866	0.0047	0.0047	21	0.679
6	250	-9.600	1.168	0.3074	0.0199	11	0.407
7	131	-19.403	2.246	0.0017	0.0030	10,14,24,25,41	0.668
8	138	-17.731	2.216	0.0039	0.0074	12,43	0.587
9	120	-8.101	1.573	0.6085	0.0407	20,24,47	0.445
10	67	-19.054	1.670	0.1844	0.0438	10,41,42,43	0.713
11	86	-15.489	2.462	0.3531	0.0631	40,42	0.746

The R-square for the model fits is fairly high for the census districts that are comprised of either a single state (Florida with $R^2 = 0.679$ and Texas with $R^2 = 0.668$) or two states (Washington. Oregon with $R^2 = 0.713$ and California. Arizona with $R^2 = 0.746$). However, census district 3 consisting of Michigan and Wisconsin has a low R-square value of 0.359. In this and other cases, state specific characteristics might be introducing some variability. For example, in census district 2 (which consists of six mid-central continental states from Ohio to Nebraska) the R-square is only 0.381. However, its model fit shows every component, including both the collapsed strata and strata within collapsed strata, to be highly significant. Neither collapsed strata nor strata in collapsed strata is significant in census district 4 (which consists of seven mid-eastern Atlantic and adjoining states), and so the low R-square value of 0.408 is perhaps indicative of the influence of state specific characteristics with respect to the occurrence of pseudo NOL farms.

Overall, the model fits are found to be useful. The total number of farms as a predictor is highly significant. Figure 4 shows a selection of typical studentized residual plots for the cases considered earlier in Figures 1 through 3.

In Figure 5, the design-based variances versus model-based variances are plotted. Except in the cases of high variance values or some potential outliers, these plots indicate reasonable model fits.

It should be pointed out that the model performance is subject to the input data and works well if the substratum variances and total number of farms are reliably estimated from the area frame data used. Otherwise, one needs to account for the errors in their estimates and investigate another modeling approach to improve upon the model performance. One possible approach that could be used to help account for the errors in the total number farms estimates used in the variance model would be to use an analogous errors in variables model.

4 SAMPLE ALLOCATION

NASS currently uses an optimum allocation algorithm for multivariate surveys as described in Bethel (1986 and 1989) to determine the sample allocation for the annual June Agricultural Survey. Linear variance constraints, which are determined by the desired level of precision in estimation for items, are specified for each item of interest, and the sample allocation that minimizes survey cost is determined. The sample

Figure 4: Studentized Residual Plots

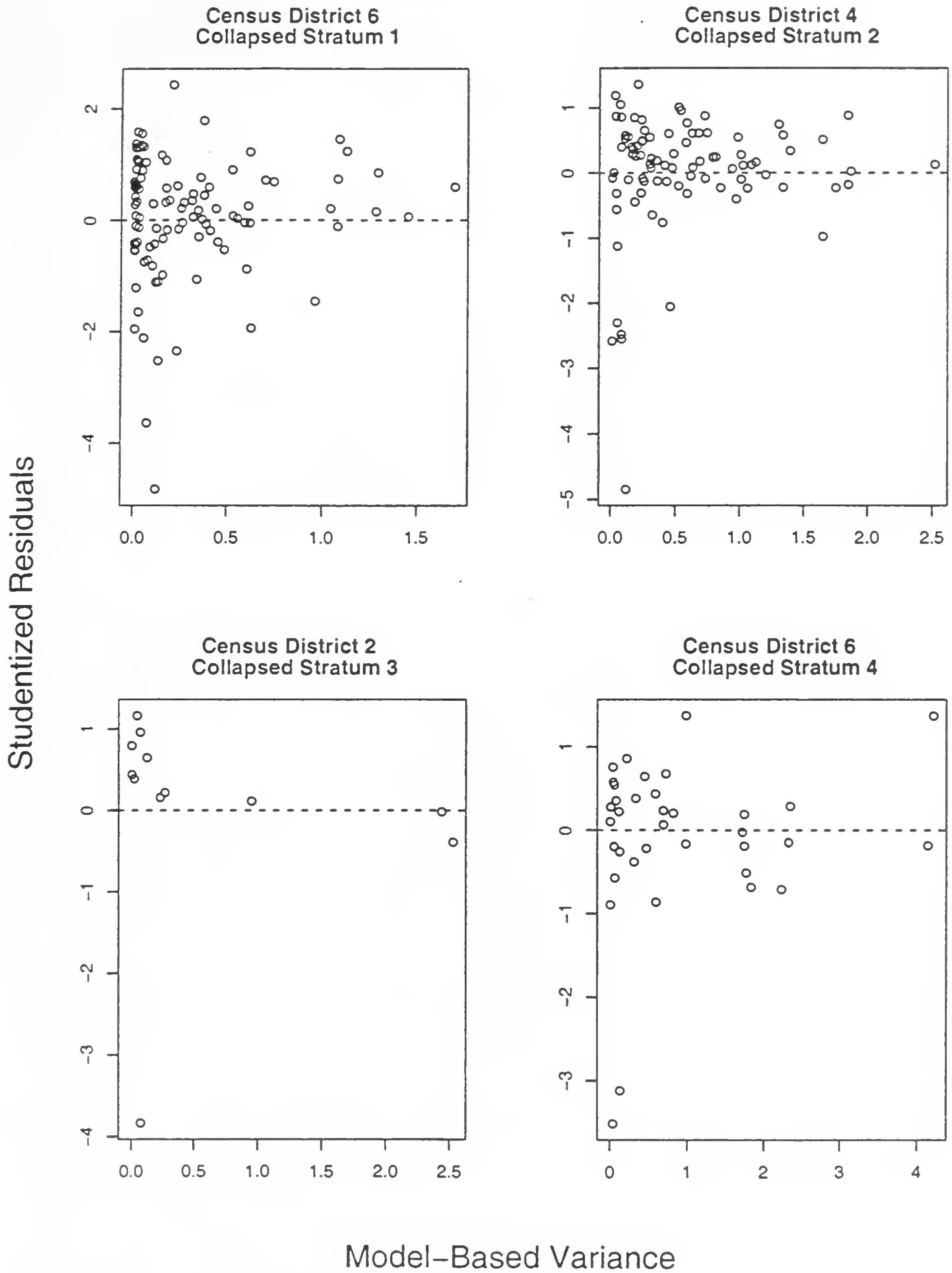
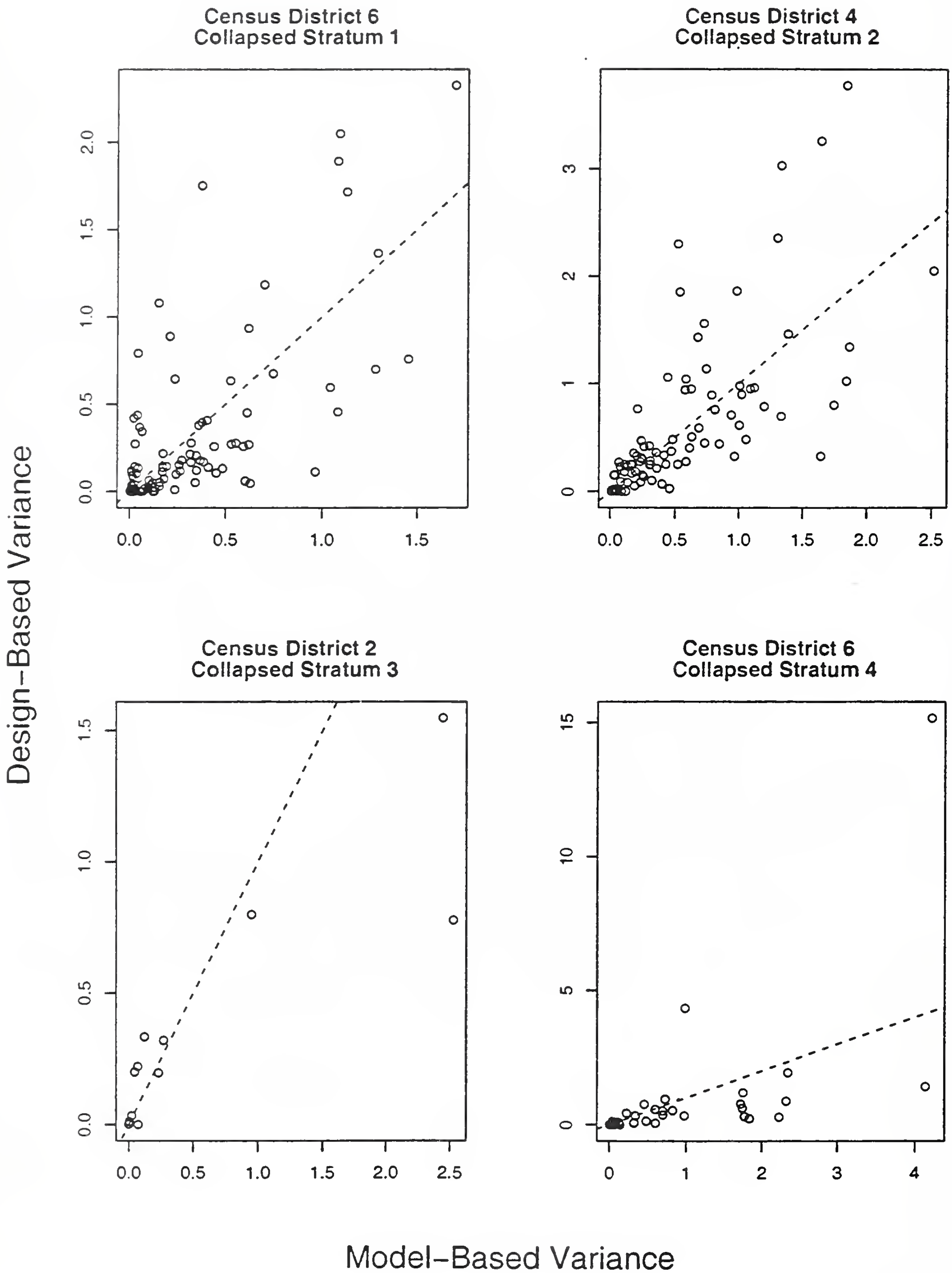


Figure 5: Model-Based versus Design-Based Variance



allocation for 1999 was based on eight agricultural items of interest: corn, spring wheat, durum wheat, soybean, cotton, winter wheat, number of NOL cattle, and number of farms. The present optimal sample allocation for estimating these items and the NML farm count at specified levels of precision is determined by adding a linear variance constraint for the case of estimating NML farm count. Several modifications are made to the optimal sample allocation to arrive at the final recommended sample allocation for the 2002 JAS. The following steps led to the final recommended sample allocation for the 2002 JAS.

1. Find the optimal sample allocation for the following cases using the allocation procedure for multivariate surveys from the 1999 JAS, including a variance constraint for NML farm count.

   Cases

   (a) Use design-based variance estimates for each stratum.

   (b) Use weighted average of design-based variance estimates within a census district to determine variance for all strata within that census district.

   (c) Use model-based variance estimates for each stratum.

   (d) Use weighted average of model-based variance estimates within a census district to determine variance for all strata within that census district.

2. Find the modified optimal sample allocation for each case (a)-(d) by setting each strata to have at least as many sample segments as in the 2002 JAS.

3. Find the R-modified sample allocation (described later in Equation 4) for the design-based and model-based scenarios. A weighted average of cases (a) and (b) above provides the design-based R-modified sample allocation. A weighted average of cases (c) and (d) above provides the model-based R-modified sample allocation.

4. Use the design-based R-modified sample allocation for those strata having large sample sizes, and the model-based R-modified sample allocation for those strata having small sample sizes.

5. Modify allocation so that each strata gets at least as many sample segments as indicated using the current NASS allocation procedure without NML farm count as an item of interest.

6. Modify allocation so that number of additional sample segments in a strata is a multiple of the number of substrata.

7. No new sample segments are added to states with allocation increases of less than 5 sample segments.

Accordingly, NASS's optimum allocation procedure for multivariate surveys is implemented with NML farm count along with the other 8 agricultural items used for the 1999 JAS. The optimal sample allocation is modified so that every stratum has at least as many sample segments as in the 2002 JAS; this is called the **modified optimal sample allocation**. The modified optimal sample allocation is determined using the design-based and model-based variance estimates. NASS currently uses design-based variance estimates in determining sample allocation. The use of model-based variance estimates can help address the problem of reliability in stratum variance estimates for NML farm count in strata with relatively few samples. The modified optimal sample allocation is also determined using the weighted average stratum variance estimates (design-based and model-based) for all strata within a census district, which are expected to be more robust estimates.

The modified optimal sample allocations are reported in Table A1 in the Appendix. Let m_h denote the modified optimal sample allocation for stratum h obtained by using the stratum specific variance estimates. Let m_H denote the sample allocation for stratum h when allocation is made using the weighted average stratum variance for all strata in a census district. Columns 7 and 8 contain m_h and m_H, respectively, using design-based variance estimates for NML farm count. Columns 10 and 11 contain m_h and m_H, respectively, using model-based variance estimates for NML farm count.

Since the sample sizes vary considerably across strata, the allocations for both the design-based and model-based variances cases are further modified to compensate for the smaller sample sizes in some strata by using a weighting technique. If there are 30 or more segments in a stratum, the allocation based on the stratum specific variance estimates is considered reliable. Otherwise, a weighted average of the allocation based on stratum specific variance estimates and the allocation based on the weighted average variance for all strata within a census district is considered. This weighted average, called the **R-modified sample allocation** for stratum h, is given by:

$$\hat{m}_h = (1 - a)m_H + am_h \tag{4}$$

where

$$a = \min(1, \frac{n_h}{\min(n_H, R)}).$$

Here n_h is the number of sample segments in stratum h in the 1999 JAS and n_H is the number of sample segments in the collapsed strata that contains stratum h in the 1999 JAS, and R is equal to the minimum number of sample segments that will support a reliable variance estimate. R is set to 30 for the purposes of this study. The R-modified sample allocation $\hat{m}_h$ is reported for the cases of design-based and model-based variances in columns 9 and 12 of Table A1, respectively.

The two R-modified sample allocations are combined to form the recommended sample allocation by using the R-modified sample allocation based on the design-based variance estimates for those strata having relatively large sample sizes (30 or more observations), and the R-modified sample allocation based on the model-based variance estimates for those strata having relatively small sample sizes (less than 30 observations). The resulting sample allocation is modified so that each stratum has at least as many samples as indicated for the 2002 JAS using the current NASS allocation procedure implemented without the addition of NML farms to the survey items of interest. It is further modified so that the additional segments are a multiple of the number of substrata. Also, if less than 5 samples were to be added to a state, then no samples were added. This is because it is not worth the data collection overhead to sample less than 5 samples in a state. As a result of this, Nebraska, North Dakota and South Dakota did not have any samples added. New England states were combined, as were Maryland and Deleware. The resulting allocation is the **recommended JAS+NML 2002 sample allocation**, which is listed in Column 5 of Table A1 along with the sampling rates given in parentheses; whereas, the JAS 2002 sample allocation is listed in Column 3 along with the sampling rates given in parentheses.

Tables A2 and A3 in the Appendix list the recommended JAS+NML 2002 sample allocation at the stratum and state level, respectively. The recommended JAS+NML 2002 sample allocation is broken down by the number of samples using the current NASS procedure (in the column labeled "Current JAS"), and the number of samples added to account for estimation of the NML (in the column labeled "Additional NML"). The recommended JAS+NML 2002 sample allocation requires an approximate 22 percent increase in sample segments over the 2002 JAS sample allocation. Table A4 in the Appendix summarizes the recommended allocation by collapsed strata at the census district level.

Table A5 in the Appendix lists the achievable SE, CV, and ratio of SE to the 1999 Board number of farms (SE/Board) for estimating the NML farm count using the 2002 recommended JAS+NML sample allocation. The CV is SE divided by the 1999 Board NOL estimate of number of farms. The SE/Board estimate is the SE divided by the 1999 Board estimate of number of farms. SE/Board of 6.5 percent or less is desirable, which requires that states be combined as follows so that this precision is achieved at a combined states level: (1) Arizona and New Mexico, (2) Delaware, Maryland and New Jersey, (3) Idaho and Oregon, (4) Nevada and Utah, (5) Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. The achievable national SE/Board is 0.5 percent for the recommended 2002 sample allocation.

5 CONCLUSIONS

Various supplemental area frame sample allocations are investigated to find one that is cost effective and leads to the desired level of reliability. Changes are recommended in area frame sample design to support both the current NASS sample survey estimates and 2002 Ag Census NML estimates in an efficient and cost-effective manner.

Estimation of stratum variances for the NML farm counts is done by modeling them as a function of the total number of farms and heterogeneity factors of stratification. Modeling provides more stable stratum variance estimates in cases when the number of sample units is small. The total number of farms is a significant predictor of the stratum variance. Farm value of sales was found to be a significant factor in modeling stratum variance, but in the final analysis, it is not included because it does not affect the recommended final sample allocation.

A modified optimum allocation is determined that reduces the achievable CV for NML farm estimates while incorporating the 2002 JAS sample allocation. An additional 2,418 sample segments are recommended for the 2002 Ag Census to significantly reduce the achievable CV from 2.8 percent using the JAS data in 1999 to a projected CV of 1.6 percent for the 2002 NML national level estimate.

Since the current area frame allocation procedure uses observed stratum variances, the resulting supplemental sample will tend to be concentrated in strata having high observed variances, which are often associated with unexpectedly large observations and/or small sample sizes. Thus, the data collected from the proposed supplemental sample can be expected to be smaller than had the data been collected from the sample determined exclusively based on the observed variances, especially when the observed variances are based on small sample sizes. The use of model-based variance estimates in the determination of the supplemental sample will tend to distribute the supplemental sample in those strata actually having high variances vs estimated high variances and thus reduce the likelihood a downward bias in the resulting estimates.

Since a multivariate allocation procedure is used, the sample allocation is affected by the estimated stratum variances of several different agricultural items. In order to achieve a more reliable and efficient sample allocation, mixed model approaches can be employed to determine more reliable stratum variance estimates for all the agricultural items included in determination of sample size and its allocation. In particular, basing the multivariate allocations on model-based variance estimates similar to those used in this report in conjunction with variance models similar to those employed by Perry (1992) would likely result in a more efficient multivariate area frame allocation.

REFERENCES

Allen, Don (1999), "1997 Census Not on Mail List Survey". U.S. Department of Agriculture, National Agricultural Statistics Service Report (unpublished).

Allen, Rich (1997), "Statistical Defensibility as Used by U.S. Department of Agriculture, National Agricultural Statistics Service," *Journal of Official Statistics*, **8**:4, 481-498.

Bethel, James (1986), "An Optimum Allocation Algorithm for Multivariate Surveys," USDA-NASS SRS Staff Report No. SF&SRB-89, January, 1986.

Bethel, James (1989), "Sample Allocation in Multivariate Surveys," *Survey Methodology*, **15**:1, 47-57.

Chhikara, Raj S. and Perry, Charles R. (1980), "Estimation of Within-Stratum Variance for Sample Allocation," Technical Report LEMSCO-14067, FC-LO-00428, JSC-16343, Houston, TX.

Chhikara, Raj S. and Perry, Charles R. (1986), "Estimation of Stratum Variance in Planning of Crop Acreage Surveys"' *Journal of Statistical Planning and Inference*, **15**, 97-114.

Fuller, Wayne A. (1987), *Measurement Error Models*, Wiley, New York.

Hallum, Cecil and Perry, Charles R. (1982), "Sampling Unit Size Considerations for Reducing Data Loads in Large Area Crop Inventorying Using Satellite-Based Data," AgRISTARS Minisymposium, JSC-18500, NASS-JSC, Houston, Texas, 1982.

Hallum, Cecil and Perry, Charles R. (1984), "Estimating Optimal Sampling Unit Sizes for Satellite Surveys," *Remote Sensing Environment*, **15**, No. 1-3.

Kott, P.S. (1990), "Mathematical Formulae for the 1989 Survey Processing System (SPS) Summary," USDA-NASS SRB Research Report No. SRB-90-08, May, 1990.

Mahalonobis, P.C. (1946), "A Sample Survey of the Acreage Under Jute in Bengal," *Sankhyā*, 4, 511-530.

Mahalonobis, P.C. (1968), "Sample Census of Area Under Jute in Bengal," *Statistical Publishing Society*, Calcutta.

National Agricultural Statistics Service (1989), *Area Frame Design for Agricultural Surveys*, Washington, D.C.: NASS, U.S. Department of Agriculture.

National Agricultural Statistics Service (1998), *1997 Agricultural Census Report*, Washington, D.C.: NASS, U.S. Department of Agriculture.

National Agricultural Statistics Service (2000), *2000 Area Frame Design Information: Implemented Year of Current NASS Area Frames*, March 20, 2000 Document, Washington, D.C.: NASS, U.S. Department of Agriculture.

Perry, Charles R. and Hallum, Cecil (1979), "Sampling Unit Size Considerations in Large Area Crop Inventorying Using Satellite-Based Data," *American Statistical Association Proceedings of Survey Research Methods*, Washington, D.C.

Perry, Charles R. and Chhikara, Raj S. (1980), "Stratum Variance Estimation for Sample Allocation in Crop Surveys," *American Statistical Association Proceedings of Survey Research Methods*, Washington D.C.

Perry, Charles R. (1992), "The Province Level Sample Size and Allocation for the Punjab, Sindh and Northwest Frontier Provinces," IPO Trip Report, NASS, USDA. Washington, D.C. December, 1992.

Perry, Charles R. (1994), "Line Segment Layouts in Terraced and Desert Areas, Data Analysis Procedures for Line Segment Data and Determination of Optimal Line Segment Lengths," Trip Report: December 31, 1994, Agricultural Data Collection Component of Agricultural Sector Support Program, U.S. Department of Agriculture, Washington, D.C.

Smith, H.F. (1936), "An Empirical Law Describing Heterogenity in the Yield of Agricultural Crops," *Journal of Agricultural Science*: **28**.

Vogel, Frederic A. (1995), "The Evolution and Development of Agricultural Statistics at the United States Department of Agriculture," *Journal of Official Statistics*, **11:2**, 161-180.

APPENDIX: Sample Allocation Tables

List of Tables

A1 Sample Allocations Based on Design-Based and Model-Based Stratum Variances 19

A2 Recommended 2002 Sample Allocation by Stratum 29

A3 Recommended 2002 Sample Allocation by State 38

A4 Recommended 2002 Sample Allocation by Census District 40

A5 Achievable CV for NML Number of Farm Estimates using Recommended 2002 Sample Allocation 41

Columns that appear in more than one table are described here.

- Columns labeled "Current JAS" contain the sample allocation for 2002 using the current NASS multivariate procedure without NML Farm Estimate as an item of interest.
- Columns labeled "Additional NML" contain the number of segments added if NML farm count is considered as an item of interest in the multivariate procedure.
- Columns labled "Recommended JAS+NML" contain the recommended allocation for 2002.

Table A1: Sample Allocations Based on Design-Based and Model-Based Stratum Variances

		Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
State	Stratum	Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
AL	13	78	(1.85%)	78	(1.85%)	78	78	78	78	78	78
AL	20	90	(0.86%)	90	(0.86%)	90	90	90	90	90	90
AL	31	4	(0.16%)	4	(0.16%)	4	5	5	4	4	4
AL	32	2	(0.19%)	2	(0.19%)	2	2	2	2	2	2
AL	40	60	(0.34%)	60	(0.34%)	60	75	60	62	74	62
AL	50	2	(4.65%)	2	(4.65%)	2	2	2	2	2	2
AZ	13	52	(2.34%)	52	(2.34%)	52	52	52	52	52	52
AZ	14	8	(2.57%)	8	(2.57%)	8	8	8	8	8	8
AZ	20	12	(1.34%)	12	(1.34%)	12	12	12	12	12	12
AZ	21	2	(2.08%)	2	(2.08%)	2	2	2	2	2	2
AZ	31	4	(0.09%)	12	(0.28%)	4	10	9	4	14	12
AZ	32	2	(0.07%)	10	(0.33%)	14	8	9	14	10	10
AZ	41	23	(100%)	23	(100%)	23	23	23	23	23	23
AZ	44	2	(0.48%)	3	(0.72%)	2	2	2	2	3	3
AZ	45	2	(0.57%)	2	(0.57%)	2	2	2	2	2	2
AZ	46	2	(0.45%)	6	(1.36%)	12	6	6	12	6	6
AZ	47	2	(0.27%)	5	(0.68%)	2	4	4	2	5	5
AZ	48	3	(0.56%)	3	(0.56%)	3	3	3	3	3	3
AZ	49	2	(0.40%)	4	(0.79%)	5	3	3	4	4	4
AZ	50	2	(0.02%)	2	(0.02%)	2	2	2	2	2	2
AR	11	312	(2.67%)	312	(2.67%)	312	312	312	312	312	312
AR	21	32	(1.18%)	32	(1.18%)	32	32	32	32	32	32
AR	31	6	(0.46%)	6	(0.46%)	6	6	6	6	6	6
AR	32	3	(0.72%)	3	(0.72%)	3	3	3	3	3	3
AR	42	70	(0.38%)	84	(0.45%)	82	88	82	96	88	96
AR	50	3	(8.57%)	3	(8.57%)	3	3	3	3	3	3
CA	11	209	(1.89%)	209	(1.89%)	209	209	209	209	209	209
CA	17	20	(0.52%)	30	(0.77%)	30	32	31	30	30	30
CA	19	6	(1.16%)	6	(1.16%)	6	6	6	6	6	6
CA	21	63	(0.94%)	63	(0.94%)	63	63	63	63	63	63
CA	27	12	(0.80%)	18	(1.20%)	26	16	20	14	15	15
CA	31	16	(0.10%)	72	(0.45%)	80	53	71	11	63	96
CA	32	2	(0.02%)	27	(0.30%)	2	22	20	2	29	27
CA	41	54	(0.47%)	108	(0.93%)	90	92	90	106	99	106
CA	45	20	(0.81%)	21	(0.85%)	22	20	21	22	20	21

NOTE:

- m_h is the modified optimal sample allocation for stratum h using stratum specific variances
- m_H is the modified optimal sample allocation for stratum h using weighted average stratum variances
- $\hat{m}_h = \alpha m_h + (1-\alpha) m_H$ is the R-modified optimal sample allocation

Continue on next page

Table A1 (continued)

		Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
State	Stratum	Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
CA	50	2	(0.51%)	2	(0.51%)	2	2	2	2	2	2
CO	13	156	(1.01%)	156	(1.01%)	156	156	156	156	156	156
CO	15	6	(0.97%)	6	(0.97%)	6	6	6	6	6	6
CO	20	35	(0.82%)	45	(1.05%)	46	44	46	52	46	51
CO	24	3	(2.78%)	3	(2.78%)	3	3	3	3	3	3
CO	25	6	(2.54%)	6	(2.54%)	6	6	6	6	6	6
CO	31	3	(0.10%)	30	(0.98%)	3	14	13	3	32	30
CO	32	3	(0.11%)	26	(0.99%)	3	12	11	3	28	26
CO	34	12	(0.94%)	16	(1.25%)	12	12	12	14	16	15
CO	35	10	(0.39%)	26	(1.03%)	13	17	16	13	29	26
CO	40	3	(0.56%)	4	(0.74%)	6	3	3	5	4	4
CO	41	2	(100%)	2	(100%)	2	2	2	2	2	2
CO	42	3	(0.57%)	3	(0.57%)	3	3	3	3	3	3
CO	43	3	(0.26%)	5	(0.44%)	3	3	3	3	5	5
CO	44	3	(0.22%)	7	(0.52%)	9	5	5	13	7	7
CO	45	3	(1.75%)	3	(1.75%)	3	3	3	3	3	3
CO	47	10	(0.21%)	21	(0.43%)	10	16	14	14	24	21
CO	48	3	(7.50%)	3	(7.50%)	3	3	3	3	3	3
CO	50	3	(0.16%)	3	(0.16%)	3	3	3	3	3	3
CT	14	2	(0.58%)	2	(0.58%)	2	2	2	2	2	2
CT	31	2	(0.05%)	6	(0.16%)	2	5	5	2	6	6
CT	40	2	(0.11%)	12	(0.66%)	8	12	12	14	11	11
CT	50	2	(16.7%)	2	(16.7%)	2	2	2	2	2	2
DE	13	10	(1.37%)	10	(1.37%)	10	10	10	10	10	10
DE	20	5	(0.52%)	6	(0.62%)	5	5	5	5	6	6
DE	31	2	(0.33%)	2	(0.33%)	2	2	2	2	2	2
DE	32	2	(3.85%)	2	(3.85%)	2	2	2	2	2	2
DE	40	2	(3.45%)	2	(3.45%)	2	2	2	2	2	2
DE	50	2	(66.7%)	2	(66.7%)	2	2	2	2	2	2
FL	13	25	(1.32%)	50	(2.65%)	52	32	49	44	30	42
FL	17	6	(0.33%)	18	(0.98%)	15	20	19	15	16	16
FL	18	2	(0.23%)	9	(1.04%)	2	9	9	2	7	7
FL	21	15	(1.38%)	15	(1.38%)	15	15	15	15	15	15
FL	22	6	(0.35%)	24	(1.41%)	27	22	23	18	18	18
FL	27	6	(0.50%)	15	(1.26%)	12	14	14	12	12	12
FL	31	6	(0.03%)	6	(0.03%)	6	6	6	6	6	6
FL	32	2	(0.01%)	2	(0.01%)	2	2	2	2	2	2
FL	40	18	(0.40%)	30	(0.67%)	21	24	22	29	31	30
FL	42	12	(0.73%)	24	(1.46%)	23	20	21	28	20	23
FL	50	2	(0.04%)	2	(0.04%)	2	2	2	2	2	2
GA	13	91	(1.56%)	91	(1.56%)	91	91	91	91	91	91
GA	20	121	(1.06%)	121	(1.06%)	121	121	121	121	121	121
GA	31	2	(0.04%)	9	(0.16%)	2	10	9	2	9	9

Continue on next page

Table A1 (continued)

		Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
State	Stratum	Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
GA	32	2	(0.22%)	2	(0.22%)	2	2	2	2	2	2
GA	40	72	(0.38%)	156	(0.82%)	158	126	158	163	126	163
GA	50	2	(2.02%)	2	(2.02%)	2	2	2	2	2	2
ID	10	30	(1.04%)	30	(1.04%)	30	30	30	30	30	30
ID	13	9	(0.86%)	9	(0.86%)	9	9	9	9	9	9
ID	15	48	(0.85%)	84	(1.49%)	84	84	84	84	84	84
ID	20	6	(1.14%)	6	(1.14%)	6	6	6	6	6	6
ID	22	16	(0.92%)	16	(0.92%)	16	16	16	16	16	16
ID	25	15	(1.15%)	15	(1.15%)	15	15	15	15	15	15
ID	31	3	(0.17%)	4	(0.23%)	4	13	12	3	4	4
ID	32	3	(0.36%)	3	(0.36%)	3	6	6	3	3	3
ID	40	3	(0.26%)	9	(0.78%)	5	5	5	5	9	9
ID	41	3	(0.27%)	8	(0.71%)	3	4	4	3	8	8
ID	42	3	(0.45%)	5	(0.74%)	3	3	3	3	5	5
ID	43	6	(1.03%)	6	(1.03%)	6	6	6	6	6	6
ID	50	3	(0.26%)	3	(0.26%)	3	3	3	3	3	3
IL	11	255	(0.87%)	255	(0.87%)	255	255	255	255	255	255
IL	12	70	(0.70%)	80	(0.80%)	73	71	73	72	72	72
IL	20	56	(0.58%)	64	(0.66%)	67	65	67	57	66	57
IL	31	12	(0.12%)	36	(0.35%)	53	26	37	52	26	36
IL	32	4	(0.03%)	20	(0.16%)	4	22	20	4	22	20
IL	33	2	(0.23%)	2	(0.23%)	5	2	2	8	2	2
IL	40	6	(0.53%)	10	(0.88%)	13	10	11	9	10	10
IL	61	2	(0.98%)	2	(0.98%)	2	2	2	2	2	2
IN	11	165	(0.93%)	165	(0.93%)	165	165	165	165	165	165
IN	12	40	(0.74%)	40	(0.74%)	40	40	40	40	40	40
IN	20	18	(0.33%)	36	(0.67%)	35	32	34	23	32	27
IN	31	25	(0.19%)	40	(0.30%)	39	42	40	38	42	39
IN	32	2	(0.03%)	9	(0.16%)	2	10	9	2	10	9
IN	33	2	(0.51%)	2	(0.51%)	2	2	2	2	2	2
IN	40	6	(0.34%)	6	(0.34%)	6	6	6	6	6	6
IN	50	2	(0.45%)	2	(0.45%)	2	2	2	2	2	2
IA	13	378	(0.87%)	378	(0.87%)	378	378	378	378	378	378
IA	20	66	(0.60%)	78	(0.71%)	75	72	75	81	73	81
IA	31	2	(0.12%)	3	(0.18%)	2	3	3	2	3	3
IA	32	2	(0.42%)	2	(0.42%)	2	2	2	2	2	2
IA	40	2	(0.28%)	4	(0.55%)	5	4	4	5	4	4
IA	50	2	(6.45%)	2	(6.45%)	2	2	2	2	2	2
KS	11	234	(0.81%)	234	(0.81%)	234	234	234	234	234	234
KS	12	112	(0.68%)	112	(0.68%)	112	112	112	128	112	128
KS	20	120	(0.48%)	150	(0.60%)	149	153	149	162	154	162
KS	31	3	(0.13%)	4	(0.18%)	3	4	4	3	4	4
KS	32	3	(0.25%)	3	(0.25%)	3	3	3	3	3	3

Continue on next page

Table A1 (continued)

State	Stratum	Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
		Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
KS	40	12	(0.45%)	14	(0.52%)	13	13	13	16	12	13
KS	50	3	(12.5%)	3	(12.5%)	3	3	3	3	3	3
KY	13	72	(0.70%)	72	(0.70%)	72	72	72	72	72	72
KY	20	60	(0.46%)	96	(0.73%)	91	84	91	84	84	84
KY	31	10	(0.13%)	20	(0.25%)	27	19	22	22	17	19
KY	32	6	(0.23%)	6	(0.23%)	6	6	6	6	6	6
KY	33	2	(0.51%)	2	(0.51%)	2	2	2	2	2	2
KY	40	27	(0.17%)	72	(0.44%)	67	81	68	82	81	82
KY	50	2	(0.48%)	2	(0.48%)	2	2	2	2	2	2
LA	13	187	(1.80%)	187	(1.80%)	187	187	187	187	187	187
LA	20	24	(0.65%)	24	(0.65%)	24	24	24	24	24	24
LA	31	4	(0.18%)	4	(0.18%)	4	4	4	4	4	4
LA	32	2	(0.15%)	2	(0.15%)	2	2	2	2	2	2
LA	40	30	(0.21%)	95	(0.67%)	95	66	95	49	65	49
LA	50	2	(3.33%)	2	(3.33%)	2	2	2	2	2	2
ME	14	12	(0.91%)	12	(0.91%)	12	12	12	12	12	12
ME	31	2	(0.31%)	2	(0.31%)	2	2	2	2	2	2
ME	40	16	(0.22%)	40	(0.54%)	29	47	37	23	41	31
ME	50	2	(10.0%)	2	(10.0%)	2	2	2	2	2	2
MD	13	15	(0.94%)	15	(0.94%)	15	15	15	15	15	15
MD	20	30	(0.60%)	36	(0.72%)	38	30	38	30	31	30
MD	31	2	(0.06%)	5	(0.15%)	2	5	5	2	5	5
MD	32	2	(0.19%)	2	(0.19%)	2	2	2	2	2	2
MD	40	10	(0.96%)	10	(0.96%)	10	10	10	10	10	10
MD	50	2	(6.45%)	2	(6.45%)	2	2	2	2	2	2
MA	14	2	(0.56%)	2	(0.56%)	3	2	2	2	2	2
MA	31	2	(0.04%)	7	(0.15%)	2	6	6	2	7	7
MA	32	2	(0.34%)	2	(0.34%)	2	2	2	2	2	2
MA	40	4	(0.13%)	22	(0.70%)	34	20	22	39	18	21
MA	50	2	(5.88%)	2	(5.88%)	2	2	2	2	2	2
MI	11	63	(0.72%)	63	(0.72%)	63	63	63	63	63	63
MI	12	24	(0.40%)	48	(0.81%)	43	36	42	27	35	29
MI	20	42	(0.39%)	77	(0.71%)	79	63	79	64	59	64
MI	31	2	(0.04%)	3	(0.06%)	2	2	2	2	3	3
MI	32	2	(0.04%)	2	(0.04%)	2	2	2	2	2	2
MI	40	10	(0.07%)	20	(0.14%)	13	15	14	23	20	21
MI	50	2	(2.74%)	2	(2.74%)	2	2	2	2	2	2
MN	11	210	(0.79%)	210	(0.79%)	210	210	210	210	210	210
MN	12	120	(0.75%)	120	(0.75%)	120	120	120	128	120	128
MN	20	35	(0.74%)	40	(0.84%)	38	35	38	39	35	39
MN	31	8	(0.11%)	20	(0.27%)	25	17	19	31	18	21
MN	32	2	(0.04%)	10	(0.20%)	2	9	9	2	11	10
MN	33	4	(0.09%)	10	(0.22%)	4	10	9	5	11	10

Continue on next page

Table A1 (continued)

State	Stratum	Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
		Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
MN	40	12	(0.16%)	30	(0.40%)	47	37	41	33	30	31
MN	50	2	(0.17%)	2	(0.17%)	2	2	2	2	2	2
MS	11	128	(2.35%)	128	(2.35%)	128	128	128	128	128	128
MS	12	40	(2.07%)	40	(2.07%)	40	40	40	40	40	40
MS	20	152	(0.81%)	152	(0.81%)	152	152	152	152	152	152
MS	31	2	(0.09%)	2	(0.09%)	2	2	2	2	2	2
MS	32	2	(0.37%)	2	(0.37%)	2	2	2	2	2	2
MS	40	30	(0.29%)	30	(0.29%)	30	45	30	30	45	30
MS	50	2	(8.70%)	2	(8.70%)	2	2	2	2	2	2
MO	11	195	(1.45%)	195	(1.45%)	195	195	195	195	195	195
MO	12	45	(0.46%)	72	(0.74%)	72	61	72	68	60	68
MO	20	45	(0.31%)	72	(0.50%)	76	70	76	104	78	104
MO	31	4	(0.11%)	4	(0.11%)	4	4	4	4	4	4
MO	32	2	(0.14%)	2	(0.14%)	2	2	2	2	2	2
MO	40	42	(0.28%)	174	(1.14%)	175	107	175	117	108	117
MO	50	2	(6.06%)	2	(6.06%)	2	2	2	2	2	2
MT	13	160	(0.63%)	170	(0.67%)	162	160	162	160	160	160
MT	20	75	(0.91%)	75	(0.91%)	75	75	75	75	75	75
MT	31	3	(0.22%)	3	(0.22%)	6	3	3	5	3	3
MT	32	3	(0.79%)	3	(0.79%)	3	3	3	3	3	3
MT	42	6	(0.94%)	6	(0.94%)	6	6	6	6	6	6
MT	43	48	(1.07%)	48	(1.07%)	48	48	48	48	48	48
MT	44	12	(0.76%)	12	(0.76%)	12	12	12	12	12	12
MT	45	3	(0.39%)	3	(0.39%)	3	3	3	3	3	3
MT	46	3	(0.42%)	3	(0.42%)	3	3	3	3	3	3
MT	50	3	(10.7%)	3	(10.7%)	3	3	3	3	3	3
NE	11	285	(0.94%)	285	(0.94%)	285	285	285	285	285	285
NE	12	77	(0.88%)	77	(0.88%)	77	77	77	77	77	77
NE	20	63	(1.32%)	63	(1.32%)	63	63	63	63	63	63
NE	31	4	(0.15%)	4	(0.15%)	7	5	5	5	4	4
NE	32	2	(0.12%)	2	(0.12%)	2	3	3	2	3	3
NE	40	40	(0.58%)	40	(0.58%)	40	45	40	40	45	40
NE	50	2	(1.09%)	2	(1.09%)	2	2	2	2	2	2
NV	13	4	(0.72%)	4	(0.72%)	4	4	4	4	4	4
NV	20	8	(1.37%)	8	(1.37%)	8	8	8	8	8	8
NV	31	2	(0.20%)	10	(1.02%)	4	5	5	6	10	10
NV	32	2	(0.43%)	5	(1.07%)	2	2	2	2	5	5
NV	41	2	(9.09%)	2	(9.09%)	2	2	2	2	2	2
NV	42	2	(0.16%)	6	(0.48%)	2	4	4	2	6	6
NV	43	2	(0.16%)	7	(0.55%)	5	5	5	10	7	7
NV	44	2	(0.15%)	6	(0.45%)	2	4	4	2	6	6
NV	50	2	(10.0%)	2	(10.0%)	2	2	2	2	2	2
NH	14	2	(1.08%)	2	(1.08%)	2	2	2	2	2	2

Continue on next page

Table A1 (continued)

		Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
State	Stratum	Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
NH	31	2	(0.41%)	2	(0.41%)	2	2	2	2	2	2
NH	40	4	(0.18%)	24	(1.10%)	45	20	23	19	19	19
NH	50	2	(14.3%)	2	(14.3%)	2	2	2	2	2	2
NJ	13	5	(1.06%)	5	(1.06%)	5	5	5	5	5	5
NJ	20	30	(1.06%)	30	(1.06%)	30	30	30	30	30	30
NJ	31	5	(0.08%)	9	(0.15%)	5	8	8	5	10	9
NJ	32	2	(0.07%)	4	(0.15%)	2	4	4	2	4	4
NJ	40	2	(0.21%)	12	(1.28%)	19	11	12	11	11	11
NJ	42	2	(0.63%)	2	(0.63%)	2	2	2	2	2	2
NJ	50	2	(6.67%)	2	(6.67%)	2	2	2	2	2	2
NM	12	6	(3.26%)	6	(3.26%)	6	6	6	6	6	6
NM	13	32	(0.98%)	32	(0.98%)	32	32	32	32	32	32
NM	20	12	(1.21%)	12	(1.21%)	12	12	12	12	12	12
NM	31	2	(0.12%)	22	(1.30%)	23	12	13	55	20	22
NM	32	2	(0.22%)	9	(1.00%)	2	4	4	2	10	9
NM	40	5	(0.33%)	9	(0.59%)	8	6	6	12	8	9
NM	41	27	(100%)	27	(100.%)	27	27	27	27	27	27
NM	42	10	(0.92%)	10	(0.92%)	10	10	10	10	10	10
NM	43	20	(1.92%)	20	(1.92%)	20	20	20	20	20	20
NM	44	2	(0.46%)	2	(0.46%)	2	2	2	2	2	2
NM	45	2	(2.22%)	2	(2.22%)	2	2	2	2	2	2
NM	46	2	(2.25%)	2	(2.25%)	2	2	2	2	2	2
NM	50	2	(0.97%)	2	(0.97%)	2	2	2	2	2	2
NY	13	25	(0.62%)	30	(0.74%)	30	32	30	33	32	33
NY	20	50	(0.35%)	60	(0.42%)	64	70	64	88	78	88
NY	31	3	(0.04%)	12	(0.14%)	3	12	11	3	13	12
NY	32	2	(0.04%)	10	(0.22%)	21	8	9	25	9	10
NY	40	12	(0.15%)	60	(0.73%)	68	57	61	61	51	55
NY	45	2	(0.12%)	9	(0.56%)	2	10	9	2	9	9
NY	50	2	(4.17%)	2	(4.17%)	2	2	2	2	2	2
NC*	13	30	(1.32%)	30	(1.32%)	30	30	30	30	30	30
NC*	20	200	(0.76%)	200	(0.76%)	200	200	200	200	200	200
NC*	31	5	(0.06%)	29	(0.33%)	25	29	26	31	26	29
NC*	32	2	(0.14%)	9	(0.64%)	2	10	9	2	9	9
NC*	40	80	(0.95%)	96	(1.14%)	93	97	93	85	97	85
NC*	50	2	(3.08%)	2	(3.08%)	2	2	2	2	2	2
ND	11	231	(0.79%)	231	(0.79%)	231	231	231	231	231	231
ND	12	90	(1.20%)	90	(1.20%)	90	90	90	90	90	90
ND	20	85	(1.06%)	85	(1.06%)	85	85	85	85	85	85
ND	31	2	(0.22%)	2	(0.22%)	2	2	2	2	2	2
ND	32	2	(0.54%)	2	(0.54%)	2	2	2	2	2	2
ND	33	2	(1.39%)	2	(1.39%)	2	2	2	2	2	2
ND	40	6	(0.26%)	6	(0.26%)	6	10	9	6	7	7

Continue on next page

Table A1 (continued)

State	Stratum	Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
		Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
ND	50	2	(1.59%)	2	(1.59%)	2	2	2	2	2	2
OH	11	110	(0.85%)	110	(0.85%)	110	110	110	110	110	110
OH	12	35	(0.60%)	49	(0.84%)	48	35	48	37	35	37
OH	20	30	(0.28%)	90	(0.85%)	88	65	88	86	66	86
OH	31	15	(0.19%)	15	(0.19%)	15	15	15	15	15	15
OH	32	5	(0.24%)	5	(0.24%)	5	5	5	5	5	5
OH	40	20	(0.22%)	40	(0.43%)	45	38	43	44	36	41
OH	50	5	(27.8%)	5	(27.8%)	5	5	5	5	5	5
OK	11	120	(1.26%)	120	(1.26%)	120	120	120	120	120	120
OK	12	36	(0.67%)	36	(0.67%)	36	36	36	36	36	36
OK	20	80	(0.60%)	80	(0.60%)	80	80	80	95	80	95
OK	31	3	(0.12%)	3	(0.12%)	3	3	3	3	3	3
OK	32	3	(0.26%)	3	(0.26%)	3	3	3	3	3	3
OK	40	90	(0.68%)	90	(0.68%)	90	90	90	119	90	105
OK	50	3	(13.6%)	3	(13.6%)	3	3	3	3	3	3
OR	10	40	(1.37%)	40	(1.37%)	40	40	40	40	40	40
OR	13	36	(0.65%)	48	(0.87%)	50	38	46	67	39	58
OR	20	36	(0.69%)	48	(0.92%)	36	36	36	43	36	38
OR	31	3	(0.05%)	21	(0.32%)	89	52	57	24	21	21
OR	32	2	(0.11%)	4	(0.23%)	2	13	12	2	4	4
OR	41	42	(0.33%)	112	(0.88%)	72	58	60	152	102	110
OR	44	10	(0.83%)	10	(0.83%)	10	10	10	10	10	10
OR	45	6	(0.11%)	39	(0.72%)	6	20	19	6	41	39
OR	50	2	(0.23%)	2	(0.23%)	2	2	2	2	2	2
PA*	13	24	(0.86%)	84	(3.00%)	84	68	84	86	69	86
PA*	20	98	(0.57%)	98	(0.57%)	86	80	86	100	90	100
PA*	31	5	(0.06%)	17	(0.21%)	24	25	24	12	28	17
PA*	32	2	(0.11%)	14	(0.77%)	2	14	12	2	16	14
PA*	40	48	(0.42%)	64	(0.56%)	40	67	60	40	59	54
PA*	50	2	(4.44%)	2	(4.44%)	2	2	2	2	2	2
RI	14	2	(3.85%)	2	(3.85%)	2	2	2	2	2	2
RI	31	2	(0.26%)	2	(0.26%)	2	2	2	2	2	2
RI	40	2	(0.49%)	3	(0.74%)	2	3	3	2	2	2
RI	50	2	(16.7%)	2	(16.7%)	2	2	2	2	2	2
SC	13	18	(1.73%)	18	(1.73%)	18	18	18	18	18	18
SC	20	60	(0.87%)	60	(0.87%)	60	60	60	60	60	60
SC	31	2	(0.05%)	7	(0.17%)	2	8	8	2	7	7
SC	32	2	(0.42%)	2	(0.42%)	2	2	2	2	2	2
SC	40	35	(0.34%)	50	(0.49%)	48	47	48	57	46	57
SC	50	2	(2.74%)	2	(2.74%)	2	2	2	2	2	2
SD	11	99	(0.76%)	99	(0.76%)	99	99	99	99	99	99
SD	12	130	(0.74%)	130	(0.74%)	130	130	130	130	130	130
SD	20	54	(0.85%)	54	(0.85%)	54	54	54	54	54	54

Continue on next page

Table A1 (continued)

State	Stratum	Current JAS Alloc	Current JAS Rate	Recommended JAS+NML Alloc	Recommended JAS+NML Rate	Design-Based Variances Case m_h	Design-Based Variances Case m_H	Design-Based Variances Case $\hat{m}_h$	Model-Based Variances Case m_h	Model-Based Variances Case m_H	Model-Based Variances Case $\hat{m}_h$
SD	31	2	(0.14%)	2	(0.14%)	6	3	3	6	3	3
SD	32	2	(0.31%)	2	(0.31%)	2	2	2	2	2	2
SD	33	2	(0.75%)	2	(0.75%)	2	2	2	2	2	2
SD	40	100	(1.34%)	100	(1.34%)	100	100	100	100	100	100
SD	44	4	(1.39%)	4	(1.39%)	4	4	4	4	4	4
SD	50	2	(0.71%)	2	(0.71%)	2	2	2	2	2	2
TN*	13	100	(0.71%)	120	(0.85%)	124	120	124	127	121	127
TN*	20	140	(0.75%)	140	(0.75%)	140	140	140	140	140	140
TN*	31	10	(0.07%)	42	(0.31%)	40	28	38	45	26	41
TN*	32	2	(0.04%)	9	(0.17%)	2	10	9	2	9	9
TN*	40	80	(0.52%)	96	(0.63%)	99	80	99	95	80	95
TN*	50	2	(0.33%)	2	(0.33%)	2	2	2	2	2	2
TX	10	110	(1.08%)	132	(1.30%)	127	123	127	120	120	120
TX	13	40	(1.19%)	40	(1.19%)	40	40	40	40	40	40
TX	14	336	(1.89%)	357	(2.00%)	343	345	343	351	346	351
TX	15	90	(1.51%)	90	(1.51%)	90	90	90	90	90	90
TX	16	20	(1.13%)	36	(2.04%)	36	36	36	37	36	37
TX	18	30	(1.06%)	42	(1.48%)	43	40	43	41	40	41
TX	20	80	(1.02%)	96	(1.22%)	92	96	92	81	89	81
TX	21	140	(1.10%)	168	(1.32%)	169	140	169	148	140	148
TX	24	84	(1.80%)	84	(1.80%)	84	86	84	84	84	84
TX	25	70	(1.47%)	84	(1.76%)	77	70	77	71	70	71
TX	26	5	(0.30%)	13	(0.79%)	5	14	13	5	12	11
TX	27	10	(0.96%)	11	(1.06%)	15	10	11	14	10	11
TX	28	25	(1.03%)	35	(1.45%)	33	35	33	30	34	31
TX	31	10	(0.06%)	10	(0.06%)	10	10	10	10	10	10
TX	32	5	(0.06%)	5	(0.06%)	5	5	5	5	5	5
TX	40	60	(0.93%)	93	(1.44%)	98	85	92	95	90	93
TX	41	56	(1.06%)	56	(1.06%)	56	68	56	58	72	58
TX	42	72	(1.05%)	76	(1.11%)	72	87	77	87	92	89
TX	43	2	(0.16%)	13	(1.02%)	2	14	13	2	15	14
TX	50	2	(2.60%)	2	(2.60%)	2	2	2	2	2	2
UT	13	28	(0.96%)	42	(1.45%)	45	29	44	28	28	28
UT	20	20	(0.80%)	20	(0.80%)	20	20	20	20	20	20
UT	31	4	(0.30%)	14	(1.06%)	5	6	6	5	14	13
UT	32	2	(0.73%)	3	(1.10%)	2	2	2	2	3	3
UT	41	5	(100%)	5	(100%)	5	5	5	5	5	5
UT	43	2	(0.25%)	4	(0.50%)	2	2	2	2	4	4
UT	44	2	(0.23%)	4	(0.46%)	2	3	3	2	4	4
UT	45	2	(0.12%)	7	(0.40%)	2	84	79	2	7	7
UT	46	2	(0.21%)	5	(0.52%)	3	3	3	12	5	5
UT	50	2	(6.06%)	2	(6.06%)	2	2	2	2	2	2
VT	14	12	(0.94%)	12	(0.94%)	12	12	12	12	12	12

Continue on next page

Table A1 (continued)

State	Stratum	Current JAS Alloc	Current JAS Rate	Recommended JAS+NML Alloc	Recommended JAS+NML Rate	Design-Based Variances Case m_h	Design-Based m_H	Design-Based $\hat{m}_h$	Model-Based Variances Case m_h	Model-Based m_H	Model-Based $\hat{m}_h$
VT	31	2	(1.25%)	2	(1.25%)	2	2	2	2	2	2
VT	40	5	(0.25%)	13	(0.65%)	10	13	13	22	12	14
VT	50	2	(22.2%)	2	(22.2%)	2	2	2	2	2	2
VA	13	4	(0.33%)	8	(0.66%)	8	7	7	5	7	7
VA	20	88	(0.28%)	154	(0.50%)	154	145	154	114	125	114
VA	31	6	(0.09%)	21	(0.32%)	38	18	22	27	18	20
VA	32	2	(0.04%)	7	(0.15%)	2	9	9	2	7	7
VA	33	2	(0.53%)	2	(0.53%)	2	2	2	2	2	2
VA	40	42	(0.21%)	84	(0.41%)	89	86	89	56	85	56
VA	50	2	(0.26%)	2	(0.26%)	2	2	2	2	2	2
WA	10	119	(1.42%)	119	(1.42%)	119	119	119	119	119	119
WA	13	60	(0.98%)	60	(0.98%)	60	60	60	60	60	60
WA	20	48	(0.61%)	48	(0.61%)	48	48	48	48	48	48
WA	31	5	(0.07%)	18	(0.23%)	5	60	42	14	20	18
WA	32	2	(0.05%)	8	(0.20%)	2	31	27	2	9	8
WA	41	27	(0.30%)	54	(0.60%)	27	35	32	34	69	57
WA	44	2	(0.69%)	2	(0.69%)	2	2	2	2	2	2
WA	45	2	(0.03%)	40	(0.69%)	2	21	20	2	43	40
WA	50	2	(0.09%)	2	(0.09%)	2	2	2	2	2	2
WV	13	10	(4.18%)	10	(4.18%)	10	10	10	10	10	10
WV	20	14	(1.58%)	14	(1.58%)	14	14	14	14	14	14
WV	31	4	(0.63%)	4	(0.63%)	4	4	4	4	4	4
WV	32	2	(2.47%)	2	(2.47%)	2	2	2	2	2	2
WV	40	34	(0.59%)	51	(0.89%)	44	47	44	34	44	34
WV	50	2	(13.3%)	2	(13.3%)	2	2	2	2	2	2
WI*	11	70	(0.91%)	70	(0.91%)	60	60	60	62	60	62
WI*	12	30	(0.33%)	60	(0.66%)	60	55	60	69	54	69
WI*	20	80	(0.51%)	96	(0.61%)	97	93	97	99	88	99
WI*	31	5	(0.05%)	13	(0.14%)	10	12	11	12	13	13
WI*	32	2	(0.04%)	2	(0.04%)	2	2	2	2	2	2
WI*	40	30	(0.32%)	30	(0.32%)	23	20	22	22	22	22
WI*	50	2	(0.50%)	2	(0.50%)	2	2	2	2	2	2
WY	11	8	(0.67%)	8	(0.67%)	8	8	8	8	8	8
WY	12	8	(0.81%)	8	(0.81%)	8	8	8	10	8	9
WY	20	8	(0.44%)	8	(0.44%)	8	8	8	10	9	9
WY	31	2	(0.25%)	8	(1.00%)	2	4	4	2	8	8
WY	32	2	(0.82%)	3	(1.23%)	2	2	2	2	3	3
WY	40	2	(0.93%)	2	(0.93%)	2	2	2	2	2	2
WY	42	5	(0.70%)	5	(0.70%)	5	5	5	5	5	5
WY	43	10	(0.96%)	10	(0.96%)	10	10	10	10	10	10
WY	44	2	(0.47%)	2	(0.47%)	2	2	2	2	2	2
WY	45	2	(0.17%)	5	(0.43%)	2	3	3	2	5	5
WY	46	2	(1.72%)	2	(1.72%)	2	2	2	2	2	2

Continue on next page

Table A1 (continued)

		Current JAS		Recommended JAS+NML		Design-Based Variances Case			Model-Based Variances Case		
State	Stratum	Alloc	Rate	Alloc	Rate	m_h	m_H	$\hat{m}_h$	m_h	m_H	$\hat{m}_h$
WY	50	2	(1.69%)	2	(1.69%)	2	2	2	2	2	2

* These states received new area frames in 2000 or 2001. The information in columns 5-12 is based on the old frame stratification and data and therefore does not reflect the current frame variance estimates or optimal allocations.

Table A2: Recommended 2002 Sample Allocation by Strata

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
4	AL	13	78	0	78
4	AL	20	90	0	90
4	AL	31	4	0	4
4	AL	32	2	0	2
4	AL	40	60	0	60
4	AL	50	2	0	2
11	AZ	13	52	0	52
11	AZ	14	8	0	8
11	AZ	20	12	0	12
11	AZ	21	2	0	2
11	AZ	31	4	8	12
11	AZ	32	2	8	10
11	AZ	41	23	0	23
11	AZ	44	2	1	3
11	AZ	45	2	0	2
11	AZ	46	2	4	6
11	AZ	47	2	3	5
11	AZ	48	3	0	3
11	AZ	49	2	2	4
11	AZ	50	2	0	2
6	AR	11	312	0	312
6	AR	21	32	0	32
6	AR	31	6	0	6
6	AR	32	3	0	3
6	AR	42	70	14	84
6	AR	50	3	0	3
11	CA	11	209	0	209
11	CA	17	20	10	30
11	CA	19	6	0	6
11	CA	21	63	0	63
11	CA	27	12	6	18
11	CA	31	16	56	72
11	CA	32	2	25	27
11	CA	41	54	54	108
11	CA	45	20	1	21
11	CA	50	2	0	2
9	CO	13	156	0	156
9	CO	15	6	0	6
9	CO	20	35	10	45
9	CO	24	3	0	3
9	CO	25	6	0	6
9	CO	31	3	27	30
9	CO	32	3	23	26

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
9	CO	34	12	4	16
9	CO	35	10	16	26
9	CO	40	3	1	4
9	CO	41	2	0	2
9	CO	42	3	0	3
9	CO	43	3	2	5
9	CO	44	3	4	7
9	CO	45	3	0	3
9	CO	47	10	11	21
9	CO	48	3	0	3
9	CO	50	3	0	3
1	CT	14	2	0	2
1	CT	31	2	4	6
1	CT	40	2	10	12
1	CT	50	2	0	2
1	DE	13	10	0	10
1	DE	20	5	1	6
1	DE	31	2	0	2
1	DE	32	2	0	2
1	DE	40	2	0	2
1	DE	50	2	0	2
5	FL	13	25	25	50
5	FL	17	6	12	18
5	FL	18	2	7	9
5	FL	21	15	0	15
5	FL	22	6	18	24
5	FL	27	6	9	15
5	FL	31	6	0	6
5	FL	32	2	0	2
5	FL	40	18	12	30
5	FL	42	12	12	24
5	FL	50	2	0	2
4	GA	13	91	0	91
4	GA	20	121	0	121
4	GA	31	2	7	9
4	GA	32	2	0	2
4	GA	40	72	84	156
4	GA	50	2	0	2
10	ID	10	30	0	30
10	ID	13	9	0	9
10	ID	15	48	36	84
10	ID	20	6	0	6
10	ID	22	16	0	16
10	ID	25	15	0	15

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
10	ID	31	3	1	4
10	ID	32	3	0	3
10	ID	40	3	6	9
10	ID	41	3	5	8
10	ID	42	3	2	5
10	ID	43	6	0	6
10	ID	50	3	0	3
2	IL	11	255	0	255
2	IL	12	70	10	80
2	IL	20	56	8	64
2	IL	31	12	24	36
2	IL	32	4	16	20
2	IL	33	2	0	2
2	IL	40	6	4	10
2	IL	61	2	0	2
2	IN	11	165	0	165
2	IN	12	40	0	40
2	IN	20	18	18	36
2	IN	31	25	15	40
2	IN	32	2	7	9
2	IN	33	2	0	2
2	IN	40	6	0	6
2	IN	50	2	0	2
2	IA	13	378	0	378
2	IA	20	66	12	78
2	IA	31	2	1	3
2	IA	32	2	0	2
2	IA	40	2	2	4
2	IA	50	2	0	2
2	KS	11	234	0	234
2	KS	12	112	0	112
2	KS	20	120	30	150
2	KS	31	3	1	4
2	KS	32	3	0	3
2	KS	40	12	2	14
2	KS	50	3	0	3
4	KY	13	72	0	72
4	KY	20	60	36	96
4	KY	31	10	10	20
4	KY	32	6	0	6
4	KY	33	2	0	2
4	KY	40	27	45	72
4	KY	50	2	0	2
6	LA	13	187	0	187

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
6	LA	20	24	0	24
6	LA	31	4	0	4
6	LA	32	2	0	2
6	LA	40	30	65	95
6	LA	50	2	0	2
1	ME	14	12	0	12
1	ME	31	2	0	2
1	ME	40	16	24	40
1	ME	50	2	0	2
1	MD	13	15	0	15
1	MD	20	30	6	36
1	MD	31	2	3	5
1	MD	32	2	0	2
1	MD	40	10	0	10
1	MD	50	2	0	2
1	MA	14	2	0	2
1	MA	31	2	5	7
1	MA	32	2	0	2
1	MA	40	4	18	22
1	MA	50	2	0	2
3	MI	11	63	0	63
3	MI	12	24	24	48
3	MI	20	42	35	77
3	MI	31	2	1	3
3	MI	32	2	0	2
3	MI	40	10	10	20
3	MI	50	2	0	2
8	MN	11	210	0	210
8	MN	12	120	0	120
8	MN	20	35	5	40
8	MN	31	8	12	20
8	MN	32	2	8	10
8	MN	33	4	6	10
8	MN	40	12	18	30
8	MN	50	2	0	2
6	MS	11	128	0	128
6	MS	12	40	0	40
6	MS	20	152	0	152
6	MS	31	2	0	2
6	MS	32	2	0	2
6	MS	40	30	0	30
6	MS	50	2	0	2
6	MO	11	195	0	195
6	MO	12	45	27	72

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
6	MO	20	45	27	72
6	MO	31	4	0	4
6	MO	32	2	0	2
6	MO	40	42	132	174
6	MO	50	2	0	2
8	MT	13	160	10	170
8	MT	20	75	0	75
8	MT	31	3	0	3
8	MT	32	3	0	3
8	MT	42	6	0	6
8	MT	43	48	0	48
8	MT	44	12	0	12
8	MT	45	3	0	3
8	MT	46	3	0	3
8	MT	50	3	0	3
2	NE	11	285	0	285
2	NE	12	77	0	77
2	NE	20	63	0	63
2	NE	31	4	0	4
2	NE	32	2	0	2
2	NE	40	40	0	40
2	NE	50	2	0	2
9	NV	13	4	0	4
9	NV	20	8	0	8
9	NV	31	2	8	10
9	NV	32	2	3	5
9	NV	41	2	0	2
9	NV	42	2	4	6
9	NV	43	2	5	7
9	NV	44	2	4	6
9	NV	50	2	0	2
1	NH	14	2	0	2
1	NH	31	2	0	2
1	NH	40	4	20	24
1	NH	50	2	0	2
1	NJ	13	5	0	5
1	NJ	20	30	0	30
1	NJ	31	5	4	9
1	NJ	32	2	2	4
1	NJ	40	2	10	12
1	NJ	42	2	0	2
1	NJ	50	2	0	2
9	NM	12	6	0	6
9	NM	13	32	0	32

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
9	NM	20	12	0	12
9	NM	31	2	20	22
9	NM	32	2	7	9
9	NM	40	5	4	9
9	NM	41	27	0	27
9	NM	42	10	0	10
9	NM	43	20	0	20
9	NM	44	2	0	2
9	NM	45	2	0	2
9	NM	46	2	0	2
9	NM	50	2	0	2
1	NY	13	25	5	30
1	NY	20	50	10	60
1	NY	31	3	9	12
1	NY	32	2	8	10
1	NY	40	12	48	60
1	NY	45	2	7	9
1	NY	50	2	0	2
4	NC	13	30	0	30
4	NC	20	200	0	200
4	NC	31	5	24	29
4	NC	32	2	7	9
4	NC	40	80	16	96
4	NC	50	2	0	2
8	ND	11	231	0	231
8	ND	12	90	0	90
8	ND	20	85	0	85
8	ND	31	2	0	2
8	ND	32	2	0	2
8	ND	33	2	0	2
8	ND	40	6	0	6
8	ND	50	2	0	2
2	OH	11	110	0	110
2	OH	12	35	14	49
2	OH	20	30	60	90
2	OH	31	15	0	15
2	OH	32	5	0	5
2	OH	40	20	20	40
2	OH	50	5	0	5
6	OK	11	120	0	120
6	OK	12	36	0	36
6	OK	20	80	0	80
6	OK	31	3	0	3
6	OK	32	3	0	3

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
6	OK	40	90	0	90
6	OK	50	3	0	3
10	OR	10	40	0	40
10	OR	13	36	12	48
10	OR	20	36	12	48
10	OR	31	3	18	21
10	OR	32	2	2	4
10	OR	41	42	70	112
10	OR	44	10	0	10
10	OR	45	6	33	39
10	OR	50	2	0	2
1	PA	13	24	60	84
1	PA	20	98	0	98
1	PA	31	5	12	17
1	PA	32	2	12	14
1	PA	40	48	16	64
1	PA	50	2	0	2
1	RI	14	2	0	2
1	RI	31	2	0	2
1	RI	40	2	1	3
1	RI	50	2	0	2
4	SC	13	18	0	18
4	SC	20	60	0	60
4	SC	31	2	5	7
4	SC	32	2	0	2
4	SC	40	35	15	50
4	SC	50	2	0	2
8	SD	11	99	0	99
8	SD	12	130	0	130
8	SD	20	54	0	54
8	SD	31	2	0	2
8	SD	32	2	0	2
8	SD	33	2	0	2
8	SD	40	100	0	100
8	SD	44	4	0	4
8	SD	50	2	0	2
4	TN	13	100	20	120
4	TN	20	140	0	140
4	TN	31	10	32	42
4	TN	32	2	7	9
4	TN	40	80	16	96
4	TN	50	2	0	2
7	TX	10	110	22	132
7	TX	13	40	0	40

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
7	TX	14	336	21	357
7	TX	15	90	0	90
7	TX	16	20	16	36
7	TX	18	30	12	42
7	TX	20	80	16	96
7	TX	21	140	28	168
7	TX	24	84	0	84
7	TX	25	70	14	84
7	TX	26	5	8	13
7	TX	27	10	1	11
7	TX	28	25	10	35
7	TX	31	10	0	10
7	TX	32	5	0	5
7	TX	40	60	33	93
7	TX	41	56	0	56
7	TX	42	72	4	76
7	TX	43	2	11	13
7	TX	50	2	0	2
9	UT	13	28	14	42
9	UT	20	20	0	20
9	UT	31	4	10	14
9	UT	32	2	1	3
9	UT	41	5	0	5
9	UT	43	2	2	4
9	UT	44	2	2	4
9	UT	45	2	5	7
9	UT	46	2	3	5
9	UT	50	2	0	2
1	VT	14	12	0	12
1	VT	31	2	0	2
1	VT	40	5	8	13
1	VT	50	2	0	2
4	VA	13	4	4	8
4	VA	20	88	66	154
4	VA	31	6	15	21
4	VA	32	2	5	7
4	VA	33	2	0	2
4	VA	40	42	42	84
4	VA	50	2	0	2
10	WA	10	119	0	119
10	WA	13	60	0	60
10	WA	20	48	0	48
10	WA	31	5	13	18
10	WA	32	2	6	8

Continue on next page

Table A2 (continued)

Census District	State	Stratum	Current JAS	Additional NML	Recommended JAS+NML
10	WA	41	27	27	54
10	WA	44	2	0	2
10	WA	45	2	38	40
10	WA	50	2	0	2
1	WV	13	10	0	10
1	WV	20	14	0	14
1	WV	31	4	0	4
1	WV	32	2	0	2
1	WV	40	34	17	51
1	WV	50	2	0	2
3	WI	11	70	0	70
3	WI	12	30	30	60
3	WI	20	80	16	96
3	WI	31	5	8	13
3	WI	32	2	0	2
3	WI	40	30	0	30
3	WI	50	2	0	2
9	WY	11	8	0	8
9	WY	12	8	0	8
9	WY	20	8	0	8
9	WY	31	2	6	8
9	WY	32	2	1	3
9	WY	40	2	0	2
9	WY	42	5	0	5
9	WY	43	10	0	10
9	WY	44	2	0	2
9	WY	45	2	3	5
9	WY	46	2	0	2
9	WY	50	2	0	2

Table A3: Recommended 2002 Sample Allocation by State

State	Current JAS	Additional NML	Recommended JAS+NML
AL	236	0	236
AZ	118	26	144
AR	426	14	440
CA	404	152	556
CO	267	98	365
CT	8	14	22
DE	23	1	24
FL	100	95	195
GA	290	91	381
ID	148	50	198
IL	407	62	469
IN	260	40	300
IA	452	15	467
KS	487	33	520
KY	179	91	270
LA	249	65	314
ME	32	24	56
MD	61	9	70
MA	12	23	35
MI	145	70	215
MN	393	49	442
MS	356	0	356
MO	335	186	521
MT	316	10	326
NE	473	0	473
NV	26	24	50
NH	10	20	30
NJ	48	16	64
NM	124	31	155
NY	96	87	183
NC	319	47	366
ND	420	0	420
OH	220	94	314
OK	335	0	335
OR	177	147	324
PA	179	100	279
RI	8	1	9
SC	119	20	139
SD	395	0	395
TN	334	75	409
TX	1247	196	1443
UT	69	37	106
VT	21	8	29

Continue on next page

Table A3 (continued)

State	Current JAS	Additional NML	Recommended JAS+NML
VA	146	132	278
WA	267	84	351
WV	66	17	83
WI	219	54	273
WY	53	10	63

Table A4: Recommended 2002 Sample Allocation by Census District

	Current JAS				Additional NML				Recommended JAS+NML			
Census	Collapsed Strata				Collapsed Strata				Collapsed Strata			
District	1	2	3	4	1	2	3	4	1	2	3	4
1	121	227	47	145	65	17	59	179	186	244	106	324
2	1761	353	83	86	24	128	64	28	1785	481	147	114
3	187	122	11	40	54	51	9	10	241	173	20	50
4	393	759	61	396	24	102	112	218	417	861	173	614
5	33	27	8	30	44	27	0	24	77	54	8	54
6	1063	333	31	262	27	27	0	211	1090	360	31	473
7	626	414	15	190	71	77	0	48	697	491	15	238
8	1040	249	32	194	10	5	26	18	1050	254	58	212
9	248	92	46	142	14	10	126	50	262	102	172	192
10	342	121	18	104	48	12	40	181	390	133	58	285
11	295	89	24	110	10	6	97	65	305	95	121	175
Total	6109	2786	376	1699	391	462	533	1032	6500	3248	909	2731

Note: Collapsed strata are defined in Table 4.

Table A5: Achievable CV for NML Number of Farm Estimates Using Recommended 2002 Sample Allocation

State(s)	1999 Number of Farms		Current	Recommended	Achievable		
	Board	NOL	JAS	JAS+NML	SE	CV	SE/Board
AL	49000	14031	236	236	1485.8	0.106	0.030
AZ, NM	23900	9967	242	299	1495.6	0.150	0.063
AZ	7900	4231	118	144	1113.2	0.263	0.141
NM	16000	5736	124	155	998.7	0.174	0.062
AR	49500	10713	426	440	1420.2	0.133	0.029
CA	89000	32537	404	556	2835.5	0.087	0.032
CO	29500	9949	267	365	1363.4	0.137	0.046
DE, MD, NJ	24800	9749	132	158	1535.6	0.158	0.062
DE	2700	1052	23	24	324.6	0.309	0.120
MD	12500	4433	61	70	1006.4	0.227	0.081
NJ	9600	4264	48	64	1113.5	0.261	0.116
FL	45000	19941	100	195	1873.5	0.094	0.042
GA	50000	15245	290	381	1748.6	0.115	0.035
ID, OR	64000	31671	363	522	3711.3	0.117	0.058
ID	24500	10377	148	198	1333.3	0.128	0.054
OR	39500	21294	177	324	3463.5	0.163	0.088
IL	79000	26042	407	469	2236.8	0.086	0.028
IN	66000	18210	260	300	1831.3	0.101	0.028
IA	97000	20937	452	467	1586.3	0.076	0.016
KS	65000	21487	487	520	1893.2	0.088	0.029
KY	90000	22632	179	270	2117.7	0.094	0.024
LA	30000	12316	249	314	1738.6	0.141	0.058
MI	52000	19668	145	215	1803.6	0.092	0.035
MN	80000	23364	393	442	2331.2	0.100	0.029
MS	42000	18011	356	356	1872.0	0.104	0.045
MO	110000	35167	335	521	2558.6	0.073	0.023
MT	27500	5844	316	326	1286.2	0.220	0.047
NE	55000	11477	473	473	1260.0	0.110	0.023
NV, UT	18000	6703	95	156	1135.6	0.169	0.063
NV	3000	1373	26	50	412.3	0.300	0.137
UT	15000	5330	69	106	1058.1	0.199	0.071
New England	27550	15595	91	181	2112.5	0.135	0.077
CT	4100	1151	8	22	443.0	0.385	0.108
ME	6900	2408	32	56	741.4	0.308	0.107

NOTE:

- "Achievable SE" column contains the achievable SE for NML Number of Farm Estimate using the recommended 2002 sample allocation
- "Achievable CV" column contains the achievable CV (SE divided by Board NOL farms)
- "Achievable SE/Board" column contains the achievable SE/Board (SE divided by Board number of farms)

Continue on next page

Table A5 (continued)

State(s)	1999 Number of Farms		Current	Recommended	Achievable		
	Board	NOL	JAS	JAS+NML	SE	CV	SE/Board
MA	6000	5513	12	35	1174.5	0.213	0.196
NH	3100	3653	10	30	1441.7	0.395	0.465
RI	750	229	8	9	167.4	0.731	0.223
VT	6700	2641	21	29	480.2	0.182	0.072
NY	38000	15469	96	183	2112.7	0.137	0.056
NC	58000	22791	319	366	2009.6	0.088	0.035
ND	31000	6145	420	423	731.2	0.119	0.024
OH	80000	30970	220	314	2315.2	0.075	0.029
OK	83000	14277	335	335	1375.6	0.096	0.017
PA	60000	24140	179	279	1993.7	0.083	0.033
SC	25000	8488	119	139	1148.5	0.135	0.046
SD	32500	7472	395	395	990.7	0.133	0.030
TN	91000	35578	334	409	2449.6	0.069	0.027
TX	226000	84262	1247	1443	3626.6	0.043	0.016
VA	49000	23309	146	278	2686.4	0.115	0.055
WA	40000	11274	267	351	1219.1	0.108	0.030
WV	21000	7102	66	83	969.1	0.136	0.046
WI	78000	24133	219	273	1893.2	0.078	0.024
WY	9200	1417	53	63	441.9	0.312	0.048
US	2185450	728083	11075	13496	11995.6	0.016	0.005